Title

Decrypting Allostery in Membrane-Bound K-Ras4B Using Complementary In Silico Approaches Based on Unbiased Molecular Dynamics Simulations.

Journal

Journal of the American Chemical Society, 146(1)

Authors

Castelli, Matteo

Marchetti, Filippo

Osuna, Sílvia

et al.

Publication Date

2024-01-10

DOI

10.1021/jacs.3c11396

Peer reviewed

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JACS
Article

Decrypting Allostery in Membrane-Bound K-Ras4B Using Complementary *In Silico* Approaches Based on Unbiased Molecular Dynamics Simulations

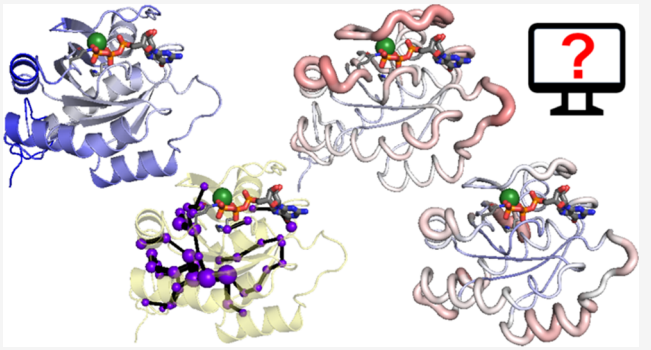
Matteo Castelli, Filippo Marchetti, Sílvia Osuna, A. Sofia F. Oliveira, Adrian J. Mulholland, Stefano A. Serapian,* and Giorgio Colombo*

Cite This: *J. Am. Chem. Soc.* 2024, 146, 901−919

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Protein functions are dynamically regulated by allostery, which enables conformational communication even between faraway residues, and expresses itself in many forms, akin to different "languages": allosteric control pathways predominating in an unperturbed protein are often unintuitively reshaped whenever biochemical perturbations arise (*e.g.*, mutations). To accurately model allostery, unbiased molecular dynamics (MD) simulations require integration with a reliable method able to, *e.g.*, detect incipient allosteric changes or likely perturbation pathways; this is because allostery can operate at longer time scales than those accessible by plain MD. Such methods are typically applied singularly, but we here argue their joint application—as a "multilingual" approach—could work significantly better. We successfully prove this through unbiased MD simulations (∼100 μs) of the widely studied, allosterically active oncotarget K-Ras4B, solvated and embedded in a phospholipid membrane, from which we decrypt allostery using four showcase "languages": *Distance Fluctuation* analysis and the *Shortest Path Map* capture allosteric hotspots at equilibrium; *Anisotropic Thermal Diffusion* and *Dynamical Non-Equilibrium MD simulations* assess perturbations upon, respectively, either superheating or hydrolyzing the GTP that oncogenically activates K-Ras4B. Chosen "languages" work synergistically, providing an articulate, mutually coherent, experimentally consistent picture of K-Ras4B allostery, whereby distinct traits emerge at equilibrium and upon GTP cleavage. At equilibrium, combined evidence confirms prominent allosteric communication from the membrane-embedded hypervariable region, through a hub comprising helix α5 and sheet β5, and up to the active site, encompassing allosteric "switches" I and II (marginally), and two proposed pockets. Upon GTP cleavage, allosteric perturbations mostly accumulate on the switches and documented interfaces.



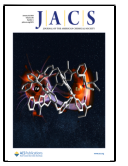
INTRODUCTION

Proteins are efficient and versatile machines that support most biochemical processes in cells.[1,2] To meet these requirements, proteins populate a diverse array of structures that are intrinsically dynamic[3−5] and are required to sustain well-defined and finely tuned functional motions.[6−8]

Allostery[6,9,10] is a fundamental mechanism regulating such functions, whereby distal parts of a protein or multimeric complex (often not intuitively linkable to active sites or binding interfaces) dynamically communicate with each other, with far-reaching repercussions on cellular activities: examples of the many aspects inextricably dependent on allostery[10] include enzyme function,[11−13] protein folding by chaperones,[14−16] signal transduction,[17] and regulation of transcription and metabolism. Understanding of allostery has been constantly expanding since seminal studies in the 1960s[18] and 1980s,[19] with the topic being extensively reviewed and reformulated.[4,7−10,20−25]

Dynamic (residue-mediated) allosteric communication pathways arise even in the more conformationally constrained proteins,[6−8,19] mediating the transition between distinct energy-wells in biomolecules characterized by well-defined "folding funnel" minima.[5,10,23,26] Such minima associated with the native structure can be rugged and different substates can exist as rapidly interconverting conformational ensembles of slightly different functional/nonfunctional forms. Allostery is the key factor regulating transitions between these forms so that they coexist in precisely the right proportions required to ensure biological functions. In this framework, a targeted allosteric

Received: October 13, 2023
Revised: December 6, 2023
Accepted: December 7, 2023
Published: December 20, 2023



https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

perturbation is what helps prompt a (physiological) change of function.[10,21,22] Such perturbation is typically introduced by either one or more specific post-translational modifications (PTMs), an endogenous ligand binding at an allosteric site, complexation with another protein, or by cleavage of a substrate. As a result, conformational equilibria are subtly altered, ushering in a "population shift",[4,7,8,10,20,21,23,26−28] as allosteric signals are relayed for a new biochemical event to occur, often far from the perturbation site(s). These events can result, *e.g.*, in the modification of the properties of a particular interface, with consequent promotion or disruption of another protein's recognition;[29,30] facilitation of conformational change;[14] substrate binding or release;[31] turnover rate of a molecular machine;[32] and regulation of enzymatic reactivity,[13] which is of particular interest in the field of biocatalysis.[8,11,12,27,31,33]

With such a delicate set of conformational equilibria required for normal biological functions, it is unsurprising that aberrant allosteric perturbations are enough to disrupt the physiological balance among different conformational populations and lead to a number of pathologies.[22] Indeed, an emerging therapeutic strategy[21,26,30] is to design small-molecule allosteric modulators[26,29,30] that bind to and interfere with identified allosteric pockets, providing a possible alternative to ineffective or toxic *ortho*steric ligands.[21,26,29]

The signal-transducing GTPase K-Ras[37,38] in its most abundant oncogenic isoform K-Ras4B[37] (Figure 1) is a textbook case of a small but allosterically complex protein (residues 2−185 when mature) consisting of a globular catalytic G domain (residues 2−166; Figure 1; yellow, purple, black, and cyan) followed by a flexible hypervariable region (HVR; Figure 1a,b; salmon) terminating with a farnesylated Cys185 that is responsible for its incorporation into the cellular membrane (Figure 1a).[35,38,39] Under healthy conditions, membrane-bound K-Ras cycles between an *active* (GTP-/Mg^{2+}-bound) state (Figure 1a,b) and an *inactive* one wherein GTP has been hydrolyzed to GDP (Figure 1c).[36] Only when K-Ras is active, key regions of its G domain (switches I and II; Figure 1)[38,40] can be allosterically remodeled to recruit and help activate various effectors,[37,38,40] which then trigger appropriate signaling cascades. Subsequent K-Ras deactivation through GTP hydrolysis also requires switches I and II to adopt specific conformations[38,40,41] and it is greatly facilitated[40,41] by the recruitment of a GTPase-activating protein (GAP), which immobilizes a catalytically crucial[41] glutamine (Gln61) in K-Ras and administers an equally crucial[41] arginine (Arg789 in GAP numbering; Figure S1b). Hydrolysis results in an inactive G domain with switches allosterically incapacitated to recruit effectors. In as many as 1 in 10 cancers,[42,43] and in common with several other oncotargets,[25] K-Ras4B is seen to undergo a plethora of different mutations[37,38,42] and/or PTMs[38] that, in ways that are not always allosterically clear[38,42] (bar steric disruption of the K-Ras4B−GAP interface, or catalytic interference, *e.g.*, through mutation of key residues Lys16 or Gln61; Figures S1b and 1),[41] hinder GTP cleavage and thus trap the protein in a harmful *hyperactive* state. In fact, while clinically relevant mutations overwhelmingly[43] concentrate in the P-loop and particularly on Gly12 and Gly13 (Figure 1; black), they can be found all over the GTPase.

In light of its relevance to oncology and unsuitability for orthosteric GTP-competitive inhibition, copious efforts have been made to identify allosteric propagation routes and hotspots in K-Ras and its mutants (*e.g.*,[37,42,44,45]), culminating in the recent (titanic) experimental effort by Weng et al.[42] to produce a

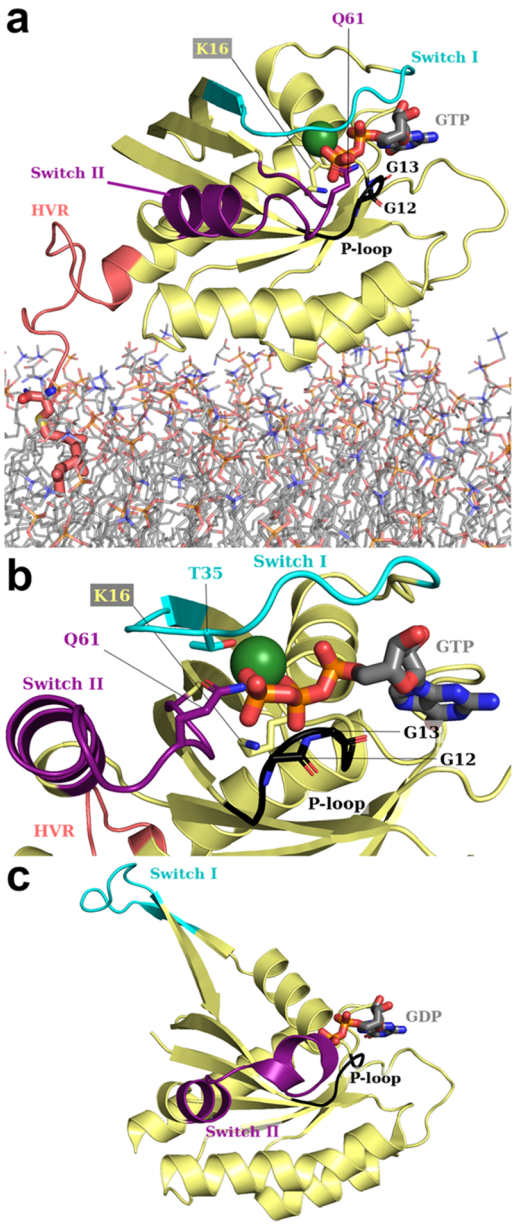


Figure 1. Structural overview of K-Ras4B, active (as simulated in this work) and inactive. (a) Zoom on K-Ras4B as it appears in our starting structure (Figure S1a). Modeled after PDB ID: 6vjj,[34] the G domain (residues 2−166; yellow) is represented with its salient features (see main text): switches I and II (in cyan and purple, respectively); and the P-loop (black; in a "reduced" definition). Catalytically relevant (Lys16, Gln61) and mutation-prone residues (Gly12, Gly13) are explicitly labeled and rendered as sticks, with C atoms and labels in the same color as their parent feature. The hypervariable region (HVR; residues 167−185; salmon) and the POPS/POPC phospholipid bilayer (lines) in which it is embedded through farnesylated Cys185 (C as large salmon spheres) are modeled from previous simulations.[35] (b) Further zoom on K-Ras4B after a ∼90° clockwise rotation perpendicular to the plane of the phospholipid bilayer (which is here omitted), revealing the same features. (c) Inactive K-Ras4B after GTP hydrolysis to GDP (PDB ID: 6mqg)[36] in the same orientation as panel (a) and showing the same features (HVR not present). In all panels, GTP (GDP in panel (c)) is represented as thicker sticks. H, Na^+, Cl^-, and solvent are omitted for clarity. Color code for explicit non-C atoms/ions: Mg^{2+} (absent from panel (c)): dark green; P: orange; O: red; N: blue; S: yellow.

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

comprehensive allosteric map of the effects of K-Ras mutations at all possible sequence positions. Significantly for this work, K-Ras4B dynamics and their variations in mutants have also been intensely studied from a computational point of view.[35,38,39,44,46−55] For a complete picture, we refer the reader to Pantsar's comprehensive review[38] but, for example, molecular dynamics (MD) simulations have confirmed: (i) that G domain switches are flexible[48,54] and flexibility is altered in different activation states[48,54] and oncogenic mutants;[51] (ii) that the G domain can fold back on the HVR and is affected by it,[46] even when K-Ras4B is embedded in the membrane;[35,47] (iii) that this interaction is weakened in oncogenic mutants;[50] (iv) that the GTPase can rotate when anchored to the membrane;[35] and (v) that there exist several allosteric pockets in the G domain[44,54] (later explicitly assessed in the aforementioned experimental allosteric map).[42] Yet, despite all of these findings and the clear medical need, covalent allosteric inhibitor Sotorasib (AMG510),[56] which rescues GTP cleavage in the oft-recurring K-Ras4B mutant of Gly12 to Cys (G12C; cf. Gly12 in Figure 1), remains the only clinically approved option so far.[42] The noncovalent pan mutant inhibitor BI-2865, impeding the restoration of the GTP-active state by preventing nucleotide exchange, was reported as late as April 2023.[57] Such paucity of medical options is indicative of the difficult druggability of K-Ras4B and the allosteric diversity of its oncogenic aberrations,[37,38,42] justifying the continued interest in this GTPase and its allosteric mechanisms.

Computational techniques[21,23,58] are essential to obtain an atomistic understanding of the determinants of allosteric regulation:[21,23−25] not only is this fundamental to drive the design of better, more "personalized" allosteric drugs and of tailored biocatalysts but also, *e.g.*,[9,25] in the field of molecular diagnostics. While the time scales accessible by atomistic simulations (now on the microsecond scale) remain on the short side when it comes to capturing complete allosteric conformational changes,[28] a number of solutions, including coarse-graining and enhanced sampling MD, have become available over the years (cf.[21,23,25,28,59]), as well as the combination with machine learning.[60] Markov State Model analysis, Replica Exchange MD, and Principal Component Analysis have also been applied specifically to K-Ras4B.[48,49,51,54]

In this context, there exist atomistic-resolution methods rooted in unbiased MD that, while starting from different initial assumptions, aim to tackle the question of unveiling the atomistic and mechanistic details of allosteric modulation and the key residues involved in this process.

Here, exploiting the considerable trove of allosteric information on K-Ras4B, we prove that a synergistic combination of four such techniques can readily interrogate 5 μs long atomistic MD simulations of K-Ras4B to provide a unique level of insight into its allosteric regulation, with each technique addressing a different aspect. Our overarching aim is to extract consensus information that will point us to the key residues/substructures and essential dynamic traits mediating allosteric regulation in K-Ras (K-Ras4B). Chosen methods to decrypt allostery include: (i) residue-pair distance fluctuation (DF) analysis, normally used to detect allosteric "crosstalk" patterns in large proteins and complexes thereof (*e.g.*,[14,16,61]); (ii) the shortest path map (SPM) method devised by Osuna and co-workers[12,31] to suggest (even distal) mutations that are most likely to allosterically impact on an enzyme's reactivity as desired; (iii) the dynamical nonequilibrium MD (D-NEMD) simulations initially developed by Ciccotti and Jacucci,[62] which map allosteric signal transduction from effector and/or substrate binding sites in receptors and enzymes;[17,63−67] and (iv) anisotropic thermal diffusion (ATD),[68] which entails heating an effector in its binding site within a (supercooled) protein or complex, and monitoring whereto the allosteric message propagates. Based on our results, we offer a view of how often seemingly independent approaches, developed from different perspectives, in reality, speak allosteric languages that are not so unintelligible and can be employed cooperatively, uncovering more information than if used on their own. We propose the concerted use of these techniques to unveil consensus mechanistic determinants and lay the basis for a more complete molecular understanding of allosteric regulation and, consequently, with more options to identify potential binding sites for novel allosteric modulators.

COMPUTATIONAL METHODS

General Procedure. To begin with, atomistic molecular dynamics (MD) simulations of membrane-embedded K-Ras4B were set up and conducted, as discussed below, in 20 independent replicas. From these, we directly derived distance fluctuation (DF)[14,16] matrices and shortest path map (SPM)[12,31] as recounted later (*i.e.*, the first two out of the four allosteric languages considered in our study). In addition, a subset of frames isolated from these equilibrium MD simulations serves as the starting point for further nonequilibrium MD simulations *per* the final two allostery detection methods (languages), *i.e.*, dynamical nonequilibrium MD (D-NEMD) simulations[17] and anisotropic thermal diffusion (ATD).[68]

System Setup. All equilibrium MD simulations were begun from a single representative structure (Figure S1a) of mature, active K-Ras4B embedded in a previously equilibrated[35] phospholipid bilayer (23% POPS, 77% POPC) through its farnesylated Cys185 (henceforth Fcy185), and solvated in an aqueous solution of 0.1 M ionic strength—achieved through the presence of appropriate amounts of Na^+ and Cl^- ions—extending on both sides of the bilayer. We prepared this structure as recounted below, ensuring that characteristics of mature K-Ras4B were modeled as closely as possible.[38]

Our template for modeling the G domain was a 1.40-Å-resolution crystal structure of wild-type (active) K-Ras4B (PDB ID: 6vjj; cf. description in Figure 1a,b).[34] From this structure, we proceeded with the *PyMol* package[69] as follows: we removed the cocrystallized domain of effector RAF1, Cl^- ions, and small-molecule ligands; the non-hydrolyzable GTP mimic GMPPNP was converted into GTP through *in situ* replacement, by an oxygen, of the nitrogen atom connecting Pγ and Pβ; and two *N*-terminal residues were stripped to obtain the mature form of the G domain (plus Lys167), complete with an *N*-acetyl cap on Thr2. To reflect its catalytically active[41] conformation, the Gln61 side chain was rotated inward to match its observed orientation in the presence of a GAP (Figure S1b; PDB ID: 1wq1);[70] all other atoms and molecules, including crystallographic waters and the Mg^{2+} cation were retained as present.

Conversely, the starting point to model the HVR in K-Ras4B and the equilibrated phospholipid membrane was a representative snapshot from one of the simulations by Prakash, Gorfe, and co-workers[35] featuring: K-Ras4B with a full farnesylated HVR region (Figure 1a; residues 167−185); a phospholipid membrane (95 POPS; 319 POPC); and aqueous NaCl solution. From this equilibrated structure, we deleted all ions and water molecules except those falling within 5 Å of any phospholipid or HVR atom, so as not to disrupt equilibrated conditions in the vicinity of the membrane and HVR. Subsequently, to merge the crystalline GTP-bound G domain and its crystallographic waters (*i.e.*, 6vjj)[34] with the equilibrated HVR, we superimposed backbone heavy atoms of residues 166−167 in both species, then deleted residues up to and including His166 in Prakash' structure, as well as residue Lys167 in 6vjj;[34] we note that the G domain in 6vjj deviates little from the simulated one (RMSD 0.785 Å). At this stage, for simplicity, we also capped the *C*-terminal Fcy185 with an *N*-methyl

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

group (rather than the *O*-methyl group that is present in the mature form).[38]

After superimposing and merging, the *reduce* utility in *AmberTools* (v. 19)[71] was employed to add hydrogens to K-Ras4B G domain residues; predict histidine tautomerization (on Nε2 in all cases and with no positively charged histidines); model optimal orientations of Asn/Gln side chains (ignoring the aforementioned catalytic Gln61); and confirm the absence of disulfide bridges. The *PropKa* package[72] predicted all residues to be in their standard protonation states at physiological pH. Finally, the *tleap* utility[71] was employed to model missing atoms in Arg73, and to (re)solvate K-Ras4B and the equilibrated membrane by reintroducing missing water molecules and randomly placing appropriate numbers of ions to neutralize the overall charge and restore the former 0.1 M ionic strength, with a final tally of one Mg^{2+}, 230 Na^+, and 135 Cl^-. The resulting system (Figure S1a) retains its original dimensions[35] (*i.e.*, a 117 × 115 × 158 Å cuboidal box), with the membrane parallel to the *xy* plane, and charged phospholipid heads about 78 and 20 Å away either box edge along the *z*-axis. The pdb file issued from *tleap*[71] was converted to .gro[73] format using the *pdb2gmx* utility: starting coordinates are available as Supporting Information.

Force Field Parameters. For reasons of mutual compatibility with other parameters, all standard K-Ras4B amino acids and terminal caps were modeled using the *ff99SB* force field[74] in its *ILDN* improvement,[75] whereas parameters for Fcy185 were based on the work by Khoury et al.[76] With regard to ions, Na^+ and Cl^- were treated with parameters by Joung and Cheatham,[77] while parameters by Allnér and co-workers[78] were used to model Mg^{2+}. For GTP and GDP (the latter present in D-NEMD simulations only; *vide infra*), we adopted the force field reported by Meagher et al.[79] Similarly, to simulate the inorganic phosphate anion $[H_2PO_4]^-$ present in D-NEMD simulations only, we introduced parameters by Kashefolgheta and Vila Verde.[80] Parameters for both lipids present in the membrane (POPC and POPS) were provided by appropriate extensions of the *Slipids* force field[81,82] by Jämbeck and Lyubartsev. Finally, the chosen water model was TIP3P.[83] Where necessary, parameters were converted to *GROMACS*-compatible[73] formats using the *acpype* code,[84] and the correct conversion of selected parameters was verified manually. Starting topologies are provided as Supporting Information; we note that despite the switch to *ff99SB-ILDN*[74,75] from force fields of the Charmm family originally used by Prakash and co-workers,[35] no major structural differences were observed in the equilibration stages of the MD simulations (cf. next subsection): this confirms that the change of force field only has a limited effect on our simulated system.

Molecular Dynamics Simulations at Equilibrium. Equilibrium MD simulations were carried out using the *GROMACS* package (version 2021.5),[73] in 20 independent 250 ns replicas (atomic velocities assigned with different random seeds). Each replica was preceded by a full steepest descent structural minimization for about 2000 steps (*i.e.*, until all atomic forces dropped below a 1000 kJ mol^{-1} nm^{-1} threshold to machine precision); and by a 200 ps equilibration stage in the *NVT* and *NpT* ensembles wherein restraints were imposed on certain atoms (details and conditions provided as Supporting Information). The 250 ns production stage for each replica, wherefore all restraints were lifted, was conducted in the *NpT* ensemble (T = 300 K; p = 1 bar), with a 2 fs time step—lengthened from the preproduction stages, see the Supporting Information—applied to the leapfrog integrator.[85] The 300 K temperature was enforced by the velocity-rescaling thermostat,[86] to which (1) protein + GTP + Mg^{2+}; (2) membrane; and (3) H_2O + Na^+ + Cl^- were coupled separately with a 100 fs time constant. To enforce the 1 bar pressure during production, we applied Berendsen's barostat[87] with a 2 ps time constant: pressure coupling was applied semi-isotropically, with the *xy* plane containing the membrane coupled separately from the *z*-axis. A 12 Å cutoff was employed to calculate the Lennard-Jones and Coulomb interactions and to determine the closest neighbors around each atom (lists were updated every 10 integrator steps). The Particle Mesh Ewald (PME) method[88] was used to calculate the Coulomb interactions, switching to reciprocal space beyond 12 Å. Lennard-Jones interactions were directly calculated up to 12 Å, and set to zero beyond this limit: effects of this were compensated, as *per GROMACS* implementation, by adding average corrections to the energy and pressure. All nonwater bonds containing hydrogens were constrained with the *LINCS* algorithm;[89] water bonds were constrained using *SETTLE*.[90] All unspecified details were set to *GROMACS* defaults;[73] all input files are provided electronically as Supporting Information.

Distance Fluctuation Analysis (DF). Distance fluctuation (DF) analysis[14,16,61] assesses whether, across an equilibrium MD trajectory or metatrajectory, all individual residue pairs in a simulated protein are moving in a more coordinated (more allosteric) or more uncoordinated (less allosteric) fashion. For a simulation of a protein composed of N residues, one thus typically obtains a single $N \times N$ **DF** matrix, whose individual elements DF_{ij} represent the average degree of coordination ("DF score") between the i^{th} and j^{th} residues. Each such element is given by the formula

$$DF_{ij} = \langle (d_{ij} - \langle d_{ij} \rangle)^2 \rangle$$

where d_{ij} represents the distance between Cα atoms of the i^{th} and j^{th} residues in a particular MD frame, and values enclosed in ⟨⟩ denote averages over a whole trajectory or concatenated trajectories (*i.e.*, $d_{ij} - \langle d_{ij} \rangle$ measures the deviation in each frame from the average distance observed throughout the simulation). There follows from this that residue pairs with a high DF score move more uncoordinatedly and are less likely to be allosterically related; *vice versa*, residue pairs exhibiting low DF scores are moving in a concerted manner and are deemed to be allosterically connected.

DF analysis on equilibrium MD simulations of K-Ras4B in this work was conducted using our *ad hoc* code,[91] directly on our 20 replicas (minus the first 5 ns of each), concatenated into a single metatrajectory; the procedure does not require any fitting or realignment.

Shortest Path Map (SPM). The theory behind shortest path maps (SPMs) is explained and justified in more detail by Osuna elsewhere:[31] the approach is based on transforming a protein into a graph of nodes-and-edges (one node = one residue), wherein edges connecting nodes i and j (the i^{th} and j^{th} residues) are only drawn if the residues' Cα atoms remain on average closer than 6 Å, and are given weighted lengths l_{ij}

$$l_{ij} = -\log(|C_{ij}|)$$

Thus, if the motion of vicinal residues i and j is more highly correlated or anticorrelated (*i.e.*, if their correlation C_{ij} approaches +1 or −1), edges connecting them will be shorter ("heavier"), signaling a greater "transfer of information" between them.

Mathematically, correlation C_{ij} between residues i and j across one or more MD trajectories can be computed as follows

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle}}$$

thus accounting for the average displacements of Cα atoms of residues i and j (Δr_i and Δr_j) from their positions in their protein's most populated conformational cluster. (As for DF equations, ⟨ ⟩ denote averages over a whole trajectory or concatenated trajectories.)

The (rather complex) node-and-edge graph is first elaborated by Osuna et al.'s *DynaComm.py* code,[12,31] which then simplifies it as follows. First, the code determines the shortest possible path (*i.e.*, along the heaviest/shortest edge or succession of edges) from each residue to every other residue: to give an idea, K-Ras4B, with its 184 residues, features 16 652 shortest paths. At the end of the process, edges most often traveled through in these residue-to-residue paths are given greater weights than those that are seldom or never passed, with the most often used edge acquiring a normalized weight of 1: a final SPM is then produced, showing residues linked by edges whose normalized weights exceed an arbitrary 0.3 threshold.

To produce a SPM, *DynaComm.py* only requires average distance matrices and correlation matrices. Conveniently, in our case, these could be directly derived from our equilibrium MD metatrajectory (*i.e.*, 20 concatenated replicas in .xtc format) using the "matrix" command in the *cpptraj* tool,[92] after a few trivial realignment steps and a clustering procedure based on Cα atoms: these steps are detailed in the Supporting Information. As Supporting Information, we also provide

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

the necessary *cpptraj* input files to perform all of these steps—including clustering and matrix derivation, the matrices themselves, and the output generated by *DynaComm.py*.

Dynamical Nonequilibrium Molecular Dynamics (D-NEMD) Simulations. Short D-NEMD simulations (6 ps perturbation + 44 ps production when surviving to completion; *vide infra* and Figure S2) were conducted starting from 71706 individual frames ("windows") directly isolated from our metatrajectory with atomic velocities (stored every 50 ps; Figure S2, thick black lines). Such frames (73.1% of the total number saved with velocities) were chosen because they were deemed to be loosely "reactive" based on the presence of a nucleophilic water molecule in the vicinity of γ-phosphate in GTP, and on Lys16 forming a hydrogen bond with one of the Oγ atoms of GTP (these criteria were inspired by a previous empirical valence bond (EVB) study,[41] and are explained in full in the Supporting Information, wherein we also provide exact "reactive" frame counts per replica).

The concept behind the D-NEMD approach, which is extensively reviewed by Oliveira et al.,[17] is illustrated in Figure S2: an identical perturbation is instantaneously introduced in each isolated window, after which MD is resumed for a short period, and the structural response of the protein is monitored over time. The structural responses at equivalent points in time are then averaged for all D-NEMD windows. In our specific case, the (chemically coherent) perturbation introduced in each window entails the immediate conversion of GTP^{4-} and nucleophilic water into GDP^{3-} and a free $[H_2PO_4]^-$ anion at their force field-dictated[79,80] equilibrium geometries. This charge-preserving interconversion is achieved by reordering and shifting the positions of as few as four atoms (details and illustration in Figure S2), without altering atomic velocities from equilibrium MD.

Starting from this new perturbed topology (provided electronically as Supporting Information), each window is first simulated for 6 ps, with an ultrashort time step of 0.6 fs, and *LINCS*[89] and *SETTLE*[90] constraints temporarily lifted (all else remains equal to the production stage of equilibrium MD; input file provided electronically). Continuing only if $[H_2PO_4]^-$ has retained its correct geometry after the initial 6 ps (the verification process is detailed in the Supporting Information),[93] we then restore the same conditions present in equilibrium MD and continue simulating each window for a further 44 ps, reaching 50 ps (Figure S2; green boxes; input file provided electronically). This ensures that deviation in Cα atoms with respect to equilibrium can be monitored in each window every 2 ps, from 6 to 50 ps. Note that no superimpositions or realignments are required prior to measuring this deviation.

Both the D-NEMD production and perturbation stages—the latter indicated as "slow growth" in previous literature,[63,64] where it was applied by gradually switching parameters between an ATP and ADP + inorganic phosphate states—come at a small loss, with as many as 80.6% of the initial "reactive" windows surviving to completion. This loss, which is due to unphysical breakup of the reconstructed $[H_2PO_4]^-$ in a minority of windows because of instantaneously shifting charges, still leaves us with a statistically sound combined D-NEMD production time of over 2.54 μs (44 ps × 57 823 windows). Exact totals are provided in Table S1. Statistical validity was confirmed by quantifying standard errors of the mean for each of the 184 Cα atom deviations, at 50 ps after hydrolysis, which revealed very small errors ranging from ±0.0028 to ±0.0053 Å (data not shown).

Anisotropic Thermal Diffusion (ATD). ATD MD simulations were conducted starting from a set of 4900 frames isolated from our 20 replicas at regular 1 ns intervals, excluding the first 5 ns of each replica. Each frame is equilibrated (supercooled) for 200 ns at 10 K, first in the *NVT* then in the *NpT* ensemble (input provided as Supporting Information), with the cutoff for Lennard-Jones (and Coulomb interactions in direct space *per* the PME method)[88] retained at 12 Å (δt = 1 fs); the neighbor list was updated every 10 integrator steps. As for equilibrium MD simulations, pressure coupling was applied semi-isotropically, with the *xy* plane containing the membrane coupled separately from the *z*-axis, at 1 bar, again using the Parrinello−Rahman barostat with a time constant of 2 ps (as in the *NpT* equilibration phase). The 10 K temperature was retained with the velocity-rescaling thermostat,[71] to which protein, membrane, and solute/solvent were coupled separately with a 100 fs time constant.

After equilibration at 10 K, we initiated 4900 production runs in the same *NpT* conditions at 10 K, except for GTP, which was instantaneously heated to 300 K (velocity-rescaling thermostat, time constant switched to 200 fs; input provided as Supporting Information),[71] and the barostat, which was switched to Berendsen's[87] as *per* the equilibrium MD production phase. The protein, remaining at 10 K, was entirely decoupled from the thermostat. Anisotropic thermal diffusion is then derived by computing per-residue RMSD of backbone heavy atoms, after structure-fitting (RMSD alignment) on backbone heavy atoms of the first frame in each ATD production run (ignoring HVR residues).

RESULTS AND DISCUSSION

We will here begin by separately presenting and discussing results from the four chosen allostery detection approaches, with only cursory references to any important similarities and differences between methods and their findings as we go along. A more systematic evaluation of consistencies and differences is provided thereafter, together with contextualization and comparison to existing computational and experimental understanding of K-Ras4B. Finally, there follows a brief critical discussion on the key implications of our findings.

Distance Fluctuation Analysis on Equilibrium MD. We previously showed and validated experimentally[94] that *higher* DF scores denote residue pairs moving in a *less* concerted manner and thus with lower allosteric dialogue; conversely, *lower* scores indicate residue pairs exhibiting *greater* allosteric coordination. We begin by analyzing the DF score matrix (Figure 2 bottom; with secondary structure elements marked along its axes) roughly in order of increasing allosteric relevance.

Analysis shows a well-defined trend whereby the unstructured HVR (residues 169−185) is moving in a generally uncoordinated fashion with respect to the G domain (cf. intense blue stripes in the top and right parts of the matrix). Other salient regions with significantly low allosteric coordination to the rest of the G domain (bluer tones) include most of switch I and the portion of switch II just downstream of catalytic[41] Gln61, including the part comprised in helix α2.

Interestingly, Thr35 on switch I and residues Thr58−Gln61 on switch II are notable outliers, with comparatively greater allosteric coupling to the G domain (whiter stripes). These outliers are all essential for reactivity: Thr58, Ala59, and Gln61 are involved in hydrolysis regulation,[41] whereas Thr35 completes the coordination shell of the GTP-chelating Mg^{2+} (and loses coordination once hydrolysis deactivates switch I).[45] Higher allosteric coupling to the bulk of the G domain is also observed for Lys16 (another catalytically relevant residue)[41] and the mutation-prone P-loop.

Finally, areas with the highest allosteric relevance in the DF matrix are spanned by regions of secondary structure, except for helix α2 and the *N*-terminus of sheet β2 (due to their partial locations on switches II and I, respectively). In particular, central β-sheets β4−β6 and helix α5 span the "whitest" areas in the matrix, suggesting they may form a coordinated and compact hub, with a high degree of interresidue crosstalk that is crucial for the repartition of allosteric signals to other K-Ras4B regions.

To aid in the interpretation of the DF matrix in Figure 2, as Supporting Information we provide a video that dynamically maps its information onto the structure of K-Ras4B: in a stepwise progression from residue 2 to residue 185, the video highlights the degree of coordination of each residue with every other residue, with projected colors on the backbone evolving

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

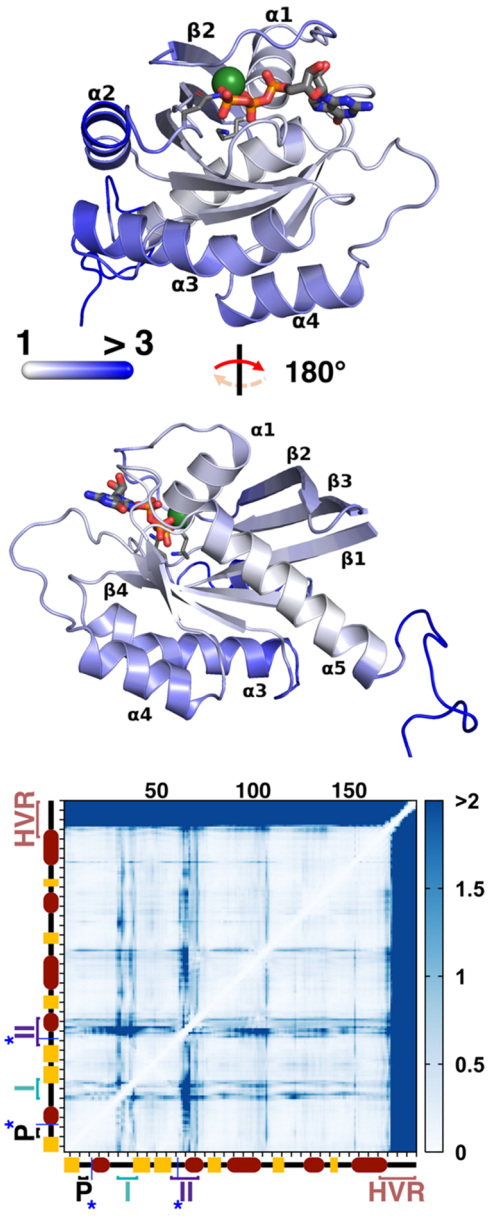


Figure 2. (top) Projection onto the starting K-Ras4B structure (Figure 1a) of the average (1D) DF score in Å^2 "felt" by each residue (see text for details), with backbone colored in increasingly darker shades of blue for increasingly higher DF scores. (middle) The same structure, rotated 180° about the axis parallel to the plane of the page. In both views: GTP is rendered as thicker sticks; K16 and Q61 as thinner sticks; Mg^{2+} as a sphere; while H, solvent, Cl^-, Na^+, and membrane are omitted for clarity; color code for atoms is identical to Figure 1a. Where possible, the secondary structure is labeled explicitly. (bottom) 2D matrix of pairwise DF scores in Å^2. Note the different scale and tones of blue. Secondary structure elements and salient conformational regions of K-Ras4B (P-loop; switches I and II; HVR) are marked along both axes for reference, as are catalytic residues K16 and Q61 (blue asterisks/lines). Tick marks along both axes are drawn every five residues, starting from the fifth residue.

accordingly. In addition to this video, in the top part of Figure 2, we have projected a 1D-averaged version of the DF matrix onto our starting K-Ras4B structure (cf. Figure 1a), which is also represented just beneath it (Figure 2 middle) after an 180° rotation. This 1D version of the matrix is obtained by summing DF scores in each matrix column (*i.e.*, residue by residue), and averaging the resulting score over the 184 pairs formable by each residue, including with itself. In doing so, one gains insight into the "average" degree of allosteric (un)coordination experienced by each residue even if the two-dimensionality of the matrix is lost. In addition to clearly recapitulating all of the allosteric traits emerging from the 2D matrix, flattening the matrix in such a way, for example, further brings to light the fact that helices $\alpha3$ and $\alpha4$ and sheets $\beta1-\beta3$ are less allosterically coordinated on average (bluer) than helix $\alpha5$ and sheets $\beta4-\beta6$ forming the central allosteric hub (Figure 2 middle structure). In addition, this operation provides per-residue DF scores that make for an easy comparison with the other allostery detection methods.

To summarize, 1D and 2D DF data concur in finding sheets $\beta4-\beta6$ to be particularly coordinated in their movements, thus exhibiting high allosteric intercommunication; the same relevance is observed for helix $\alpha5$. Most of the G domain only retains a moderate degree of internal coordination, with the P-loop; possibly sheets $\beta1$, $\beta3$, and most of $\beta2$; catalytic Lys16 as part of helix $\alpha1$; and (to a lesser extent) Gln61 and very minor portions of switches I and II falling in this category. In contrast, the HVR and the greater part of effector switches I and II are found to have a very low degree of allosteric coupling with other regions, exhibiting clear uncoordinated movements.

SPM Derived from Equilibrium MD. The shortest path map (SPM) representing the main allosteric communication route[12,31] across K-Ras4B, as derived from our full metatrajectory, is illustrated in Figure 3 as purple spheres (residues) connected by black lines (paths). The *PyMOL* session file used to derive Figure 3 (left) is also provided, along with all SPM output and input, as Supporting Information. As a reminder, the (dimensionless) "shortness" of each path segment is proportional to its thickness in Figure 3, and proportional to the intensity of allosteric communication between the two vicinal residues it connects.

The SPM paints a very similar picture to the DF analysis, as commented later in this subsection. Indeed, the SPM features a prominent allosteric pathway that originates midway along the HVR (Figure 3 bottom)—close to the membrane anchor point and best describable in the *C*- to *N*- direction—that eventually branches out to reach most allosterically relevant G domain regions, while virtually avoiding (most of) switches I and II. Working its way up the HVR and the *C*-terminal half of the $\alpha5$ helix, the path first reaches Ile163; Ile163 communicates very strongly with residues Val112, Met111 at the *N*-terminus of sheet $\beta5$, and Pro110 on the $\alpha3-\beta5$ loop, with an elevated relative shortness of 0.79 out of 1 (Figure 3 bottom; *cf.* Computational Methods). Indeed, by virtue of this shortness, Ile163 and the three $\alpha3-\beta5/\beta5$ residues are seen to form the main allosteric communication hub in K-Ras4B: in particular, the path connecting Val112 and Met111 themselves is the shortest one in the protein.

Five shortest path branches depart from this four-residue hub. We label these I–V and discuss them in detail in the Supporting Information. From an allosteric point of view, Branch V is the most intensely traveled route and the one spanning the most interesting regions. It branches off to $\beta4$ from the hub on $\beta5$, with a dual coupling from Pro110 ($\beta5$) to Gly77 ($\beta4$) and from Met111 to Phe78 ($\beta4$); from there, it further branches out into three subbranches V.1, V.2, and V.3 (cf. Supporting Information), of which V.3 is the most allosterically interesting, since it reaches all crucial areas of the active site. More specifically, moving through Val7 on $\beta1$ and the *C*-terminal residues of $\beta3$, it reaches all the way up to residues Thr58 and

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

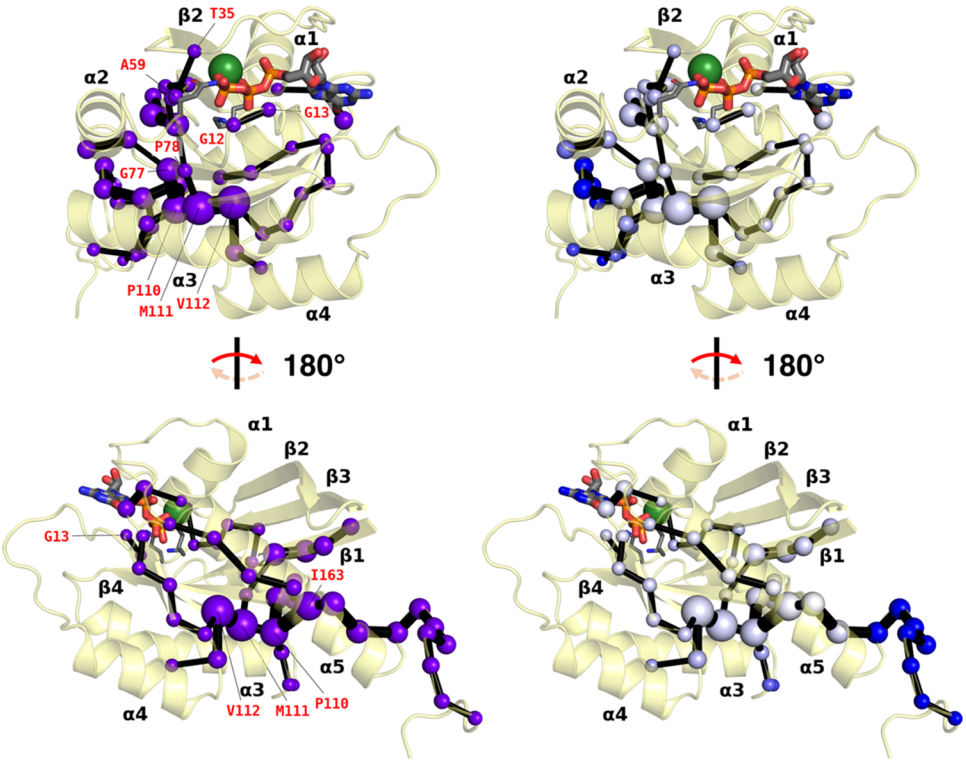


Figure 3. Allosteric shortest path map (SPM) calculated from our 20 concatenated equilibrium MD simulations of K-Ras4B. On the left, we show the simple SPM drawn on the starting structure: purple spheres represent communicating residues (Cα atoms) and the thickness of black sticks is proportional to their "allosteric closeness" (see text); where possible, residues mentioned in the main text are labeled in red. On the right, we color the SPM according to the 1D DF score of residues that are present (cf. Figure 2 top panels). Views and atom color codes are identical to those for Figure 2, with added transparency.

Ala59 on switch II, which as mentioned during DF analysis are among residues controlling hydrolysis.[41] Equally intriguingly and again in line with DF analysis, V.3 also encompasses Thr35, which keeps Mg^{2+} coordination; as well as the oft-mutating Gly12 and Gly13 in the P-loop. In fact, the importance of some of the most crucial residues along the SPM is also confirmed by mutagenesis studies:[42] we comment on this in more depth later on in this section.

In short, the shortest path originates from the HVR: this likely picks up allosteric messages from the membrane and conveys them to a major hub located on helix α5, loop α3–β5, and the *N*-terminus of sheet β5. From this hub, the path branches out to fully encompass sheets β1, β4, and β6. Crucially, key areas within the binding site or its vicinity are also eventually reached, most notably Thr58 and Ala59 on switch II; Thr35 on switch I; and the P-loop. On the other hand, regions such as helix α3, the remainder of switches I and II (minus Thr58 and Ala59), and the remainder of sheet β2 remain out of reach.

Indeed, consistency of SPM findings with DF data is readily ascertainable (Figure 3 right), particularly in terms of the 1D average DF scores projected on the structure in the top and middle of Figure 2. With the sole exception of the HVR, which only the SPM identifies as allosterically relevant, all areas touched by the SPM are also found to be allosterically important by our DF analysis, both in terms of mutual allosteric crosstalk and in terms of (1D) average coordination (Figure 2). This is clearly demonstrated by the fact that blue spheres on the right of Figure 3 (*i.e.*, high average DF scores, poor allosteric coupling) are clearly limited to HVR residues. *Vice versa*, all remaining areas with low average coordination coincide with those untouched by the SPM.

D-NEMD Simulations. D-NEMD simulations do not assess allostery under equilibrium, unlike SPM or DF analysis, but they instead model its role in relaying information away from the nucleotide binding site upon forced GTP hydrolysis to GDP and $[H_2PO_4]^-$. When mapped onto the initial K-Ras4B structure, for example, 36 ps after hydrolysis (Figure 4), the average structural response of the protein reveals a very interesting hydrolysis propagation pattern. We should note that the pattern's progression is uniform throughout the monitored 50 ps hydrolysis period: to follow it in full from 0 to 50 ps, the reader is referred to the video we provide as Supporting Information.

In any case, areas in the immediate proximity of the binding site evidently feel the greatest effects from hydrolysis (redder color = stronger deviation in Figure 4), but not in a uniform way: it is clear (Figure 4 top) that switches I and II in their entirety bear a greater brunt. In fact, with a 2.27 Å deviation at 36 ps, Glu63 on switch II is the residue with the greatest average deviation compared to equilibrium MD, and catalytic Gln61 itself deviates by 2.01 Å. By comparison, less deviation is observed in other areas adjacent to the GTP binding site, namely, the β6–α5 loop and the P-loop, whose average deviations of 1.79 and 1.74 Å, respectively, make them appear much whiter in Figure 4 (top; the β6–α5 loop appears behind the guanine moiety of GTP). Catalytic residue Lys16, despite its proximity to the cleaving phosphate, is basically unaffected in the first 50 ps after hydrolysis.

Compared to equilibrium MD, sheet β1, core sheets β4–β6, and helix α5 (Figure 4 bottom) exhibit an even lower combined average deviation at 36 ps of 1.48 Å: this echoes their rigidity in SPM and DF findings and suggests that while they could act as rigid transit hubs for the transmission of allosteric signals, they

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901–919

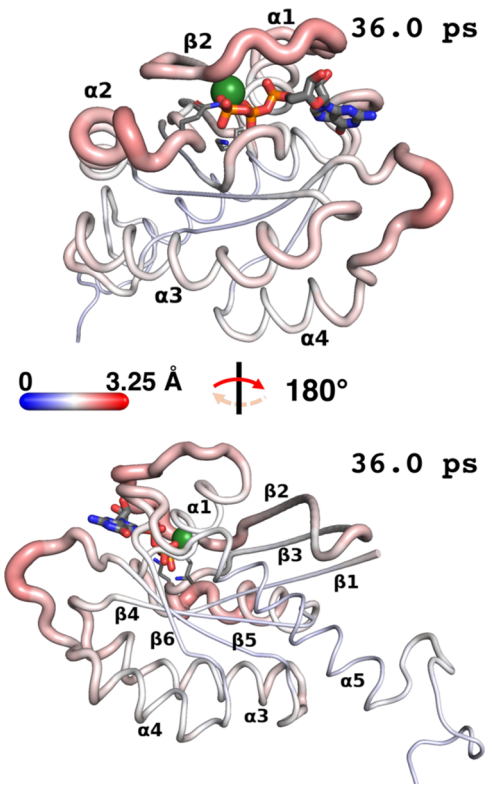


Figure 4. D-NEMD simulation results viewed on the starting K-Ras4B structure (views and atom color codes are identical to those for Figure 2). Backbone thickness of each residue and intensity of red vs blue is proportional to the average direct deviation (in Å) of its Cα atom in D-NEMD simulations—*i.e.*, after forced GTP hydrolysis—from its position in equilibrium MD simulations at an equivalent moment in time. Deviations are averaged over 57823 "reactive frames" surviving hydrolysis (Table S1); in this figure, they are exemplified as they appear 36.0 ps after hydrolysis, but they were measured up to 50.0 ps after (see video in the Supporting Information). The choice of time is the one that visually maximizes contrast between most- and least-deviating residues at the chosen color scale (0−3.25 Å deviation); note, however, that no residue reaches either of these values at the chosen time.

are themselves unaffected by hydrolysis. Focusing on the remaining peripheral regions in K-Ras4B, it is equally clear from Figure 4 that not all of them are affected by GTP hydrolysis in the same way: as in previous D-NEMD studies,[63−66] greater flexibility and/or solvent exposure is by no means synonymous with a greater deviation upon perturbation and, indeed, eliminating most deviational "noise" that is not directly linked to the introduced perturbation is precisely one of the specific advantages of the D-NEMD approach. The only distal region from the hydrolyzing phosphate to undergo significant deviation, though still part of the binding site, is evidently the $\beta5-\alpha4$ loop (on the right of Figure 4 top; and on the left of Figure 4 bottom), on par with deviations observed for residues in switch II; remaining peripheral regions do experience some of the effects of hydrolysis after 36 ps (cf. paler red areas in Figure 4), but not to the same magnitude; these include, *e.g.*, the $\alpha1$ helix ($\lesssim$1.98 Å) in its farthest part from GTP, the $\beta2-\beta3$ loop ($\lesssim$1.95 Å), the *N*- and *C*-termini of helix $\alpha3$ ($\lesssim$1.89 Å), the *N*-terminus of helix $\alpha4$ ($\lesssim$1.87 Å), and the $\beta2$ sheet outside switch I ($\lesssim$1.79 Å). Deemed allosterically important by the SPM approach but recognized by DF analysis as one of the K-Ras4B regions with the greatest flexibility and solvent exposure, the (very peripheral) HVR stands out for its particularly low deviation ($\lesssim$1.64 Å), entirely comparable to the $\alpha5/\beta1/\beta4-\beta6$ core: this again suggests it is likely unaffected by allosteric signals emanating from the active site.

The reader will recall (*cf.* Computational Methods) that D-NEMD statistics were collected on 57823 loosely "reactive" windows surviving hydrolysis. In defining these, we purposely ignored the conformation of Gln61, which in reality is crucial to properly lock the nucleophile into position for attack,[41] since it would have lowered the number of viable frames. Still, we can confirm that if, for completeness, perturbations are recalculated on this fraction of surviving D-NEMD frames in which Gln61 too is suitably positioned for hydrolysis (23209; criteria in the Supporting Information), the result is virtually indistinguishable from Figure 4 (data not shown but available upon request).

To recapitulate, D-NEMD simulation data show that the effects of the forced hydrolysis of GTP have immediate and significant repercussions on switches I and II. Other areas in the vicinity of the nucleotide either feel these effects considerably less, including the P-loop and the $\beta6-\alpha5$ loop, while catalytic Lys16 is barely affected. Away from the active site, the rigid $\alpha5/\beta1/\beta4-\beta6$ core and the HVR are largely unaffected by hydrolysis; whereas some of its effects do indeed propagate to other peripheral regions, most prominently to the $\beta5-\alpha4$ loop.

ATD Simulations. Like D-NEMD simulations, anisotropic thermal diffusion (ATD) simulations too assess propagation of allosteric crosstalk patterns upon perturbing our system from equilibrium. In this case, however, the perturbation does not consist in forced GTP hydrolysis, but in reheating GTP alone to 300 K after supercooling the entire system at 10 K, as repeated starting from 4900 individual frames isolated from the MD metatrajectory. Allosteric signals irradiating from the binding site in ATD simulations are, therefore, conceptually different compared to those in D-NEMD simulations: they should capture areas of K-Ras4B that are allosterically dialoguing with GTP as a whole rather than those sensing the effects of GTP hydrolysis. Also, unlike the D-NEMD approach, progression of allostery in ATD is measured with respect to the first production frame in each window (when all atoms will have shifted by some degree), *not* with respect to an exactly equivalent time at equilibrium (when shifts are only concentrated in perturbed areas): this means that noise resulting from ordinary K-Ras4B flexibility is not entirely canceled. Quite on the contrary, we found that including the HVR when aligning to the first frame of an individual window always led to generally high levels of "RMSD noise" across the K-Ras4B structure, making it hard to extract meaningful allosteric information. For this reason, we proceeded to exclude most of the HVR (residue 169 and above) from alignment to the first frame in each ATD window and exclude it from the ensuing ATD analysis (Figure 5).

ATD simulation data projected on the starting K-Ras4B structure (Figure 5) show that while areas of greater average RMSD (redder) with respect to the start of GTP heating generally coincide with areas of greater absolute deviation in D-NEMD simulations after hydrolysis (Figure 4), the extents by which they deviate in the two situations can be quite distinct. The most patent difference is the more moderate deviation in switches I and II compared to other areas: this is both in contrast to D-NEMD simulations, whereby they are the part of K-Ras4B that is most affected by hydrolysis (Figure 4), and in contrast to their uncoordinated (flexible) status detected in the SPM (Figure 3) and by DF analysis (Figure 2). More concretely, the maximum average RMSD detected upon reheating GTP is 0.22 Å (Asp30) in switch I and 0.23 Å (Ala66) in switch II/helix $\alpha2$

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

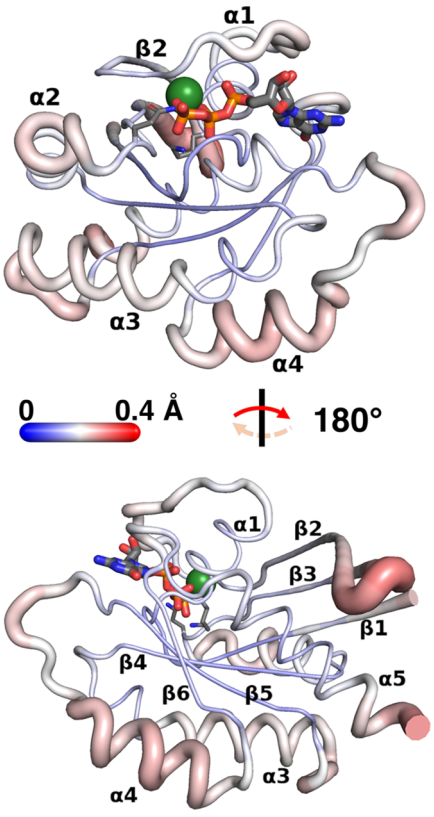


Figure 5. ATD results viewed on the starting K-Ras4B structure (views and atom color codes are identical to those for Figure 2; note, however, that in this case, analysis was truncated beyond residue 168, and therefore most of the HVR is omitted). Backbone thickness of each residue and intensity of red vs blue is proportional to the average RMSD (in Å) reached by that residue by the end of an ATD production run compared to its beginning. Averages are taken over the 4900 individual frames on which we performed ATD simulations.

(Figure 5 top): even visually, it is clear that there are areas in K-Ras4B that *per* Figure 5 show greater or comparable average deviation, all of which are far from the GRP binding site. The first of these is $\beta2-\beta3$ loop (Figure 5 bottom), whose Gly48 exhibits the greatest deviation of all residues considered (0.29 Å), followed by helix $\alpha4$, with an average deviation of 0.24 Å. The $\alpha3-\beta5$ loop, whose residue 107 incidentally stands out as allosterically uncoupled in the DF analysis (Figure 2), shows a comparable degree of average deviation (Figure 5 bottom), and the $\beta5-\alpha4$ loop deviates just a little less, at 0.18 Å on average (Figure 5 top and bottom).

The only other region of the active site that shows appreciable deviation—which, incidentally, is found by the D-NEMD simulations too—is the $\beta6-\alpha5$ loop, also at 0.18 Å (Figure 5 top; behind the guanine moiety in GTP). Conversely, what stands out in the active site is the absence of significant deviation in the P-loop (Figure 5 top), and in catalytic residues Gln61 (Figure 5 top) and Lys16 (Figure 5 bottom), suggesting that while coupled to GTP once it is hydrolyzing (D-NEMD), these areas are not relevantly coupled to GTP prior to it.

Focusing, finally, on the least-deviating parts of K-Ras4B away from the active site (Figure 5), we once again confirm that the $\alpha5/\beta1/\beta4-\beta6$ core is minimally perturbed, excluding the final few C-terminal residues of $\alpha5$, which begin to feel the strong deviation experienced by the rest of the (unincluded) HVR group.

In summary, areas that are perturbed upon heating GTP in ATD simulations qualitatively overlap with areas feeling the effects of GTP hydrolysis in D-NEMD simulations. However, it is interesting to note that all perturbed areas distal to the binding site experience a greater average deviation compared to those in its proximity, including switches I and II. The P-loop and catalytic residues Gln61 and Lys16 are significant outliers, showing no allosteric coupling to GTP heating at all; the central $\alpha5/\beta1/\beta4-\beta6$ core is also found to be unaffected by GTP heating, but in this case much expectedly.

Listening to the Various Languages. While clear trends already emerge from isolated analyses in Figures 2−5, to obtain meaningful information about K-Ras4B, it is of course necessary to compare the four allostery detection methods more systematically: this is possible through the heatmap plotted in Figure 6. In the heatmap, which only spans residues 2−168 since

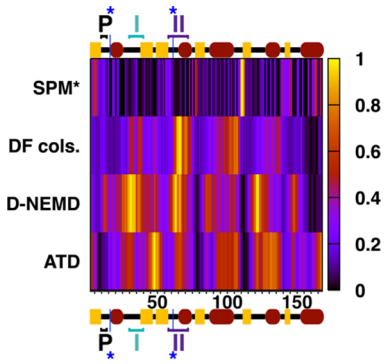


Figure 6. Combined overview of the four allosteric languages compared in this work, with per-residue representative scores for each method (see main text for chosen ones) normalized to between 0 (black) and 1 (yellow), passing through shades of purple, red, and orange. Since the ATD analysis excludes the HVR, all comparisons have been truncated to span residues 2−168 only: secondary structure elements, the P-loop, and switches I and II have been marked along the top and bottom axes, with the two blue (*) denoting positions of catalytically relevant Lys16 and Gln61. *The chosen SPM score (see main text) denotes greater residue rigidity when closer to 1 and is presented in the top row to distinguish it from other method scores, which indicate more rigidity when closer to 0.

ATD was not meaningfully analyzable in the HVR, we compare a distinctive raw per-residue score S_{raw} chosen for each allosteric language, rescaled/normalized so that the score $S_{norm,i}$ for the i^{th} residue will always fall between 0 (black) and 1 (yellow).

Full details about derivation of S_{norm} scores from S_{raw} are provided as Supporting Information. Regarding the process of choosing a suitable per-residue S_{raw} for each method, ATD and D-NEMD simulations (Figure 6 two bottom rows) already provide per-residue scores (respectively, RMS backbone deviation and Cα deviation from equilibrium). S_{raw} for DF analysis (Figure 6 second row) are simply taken to be average per-residue DF scores emerging from the flattening of the 2D matrix (*i.e.*, those projected on the structure in Figure 2). For SPM (top row of Figure 6), we choose a S_{raw} that is derivable from the *DynaComm.py* code (SPM prominence) and is proportional to the number of times that a given residue is "visited" as an intermediate point along one of the many individual *shortest paths* that connect each residue to every other residue.

Of note, per-residue S_{norm} scores for SPM are conceptually opposite to the other three chosen scores. In the former case,

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

residues with the highest S_{norm} correspond by definition to those falling along the final SPM (Figure 3): residues with scores tending to 1 represent, therefore, those with the highest degree of allosteric communication. Conversely, in the other three languages, scores tending to 1 either denote residues with *low* allosteric communication/rigidity or those with the greatest deviation upon heating or hydrolysis. The top row of Figure 6, therefore, follows an opposite chromatic trend compared to the bottom three.

Interpreting the Various Languages and Speaking the Languages of Experiment. With this *caveat* in mind, Figure 6 unequivocally indicates that there is a generally high degree of consensus between all four approaches in categorizing several important regions of K-Ras4B. At the same time, consensus is clearly not universal, and there are some differences between certain languages: these differences are, incidentally, entirely expectable since they directly reflect the fact that, as set out in the *Introduction*, not all allosteric routes are created equal.[6] More specifically, there will be certain pathways that a system will preferentially be prone to explore only when at equilibrium by virtue of its intrinsic dynamics; there will be other pathways whose exploration will only be able to gather some pace after a biochemical trigger (in our case, GTP hydrolysis); and there will be allosteric pathways that remain significant in both circumstances, possibly to different extents.

It is precisely due to the different declensions of allostery that we here invoke the use of more than one language in the first place, and that, in general, so very diverse allostery detection methods have emerged over the years.[21,25] A full portrait of the complex allostery of K-Ras4B in which the role of each residue is meticulously reconstructed is only possible, as we shall see, thanks to the unique nuances that each allostery detection method is able to provide, interrogating K-Ras4B on its propensity to visit allosteric states that become relevant at very different stages of its biochemical lifecycle.

Broadly speaking, the SPM and the DF analysis are natural partners (akin to linguistic cognates), requiring no additional information other than the original set of unbiased MD simulations from which they are constructed. In our specific case, they inform us from slightly different perspectives on likely allosteric hotspots characterizing the GTP-active state. Additionally, by virtue of their two-dimensional nature (Figure 2 bottom), DF scores also allow one to break down general allosteric signals into components of "individual" allosteric dialogue between specific residues or groups of residues. Even if our system does not undergo major conformational rearrangements during MD, the SPM and DF analysis uncover allosteric hotspots likely to mediate such rearrangements at sufficiently long time scales and those that would be most disruptive if interfered with.

ATD and D-NEMD simulations also revolve around the same set of unbiased MD simulations but, each in its own way, they inform us on allostery by monitoring the average *change* in dynamics upon introduction of a specific perturbative event. They identify "temporary hotspots" that, at a given point in time, will harbor the greatest repercussions form the allosteric perturbation in question. Allosteric propagation routes (responsive residues) can in principle be indirectly deduced by retracing the perturbation backward or forward in time, and thus distinguished from regions whose dynamics are unaffected simply because they are not involved in conveying that particular perturbation. By artificially heating GTP in a supercooled K-Ras4B, the more "unphysical" ATD reveals allosteric signals specifically emanating from the active site. D-NEMD simulations similarly inform explicitly on temporary allosteric hotspots at (a) given time(s) after perturbation, but is based on instantaneous differences resulting from a real biochemical event (phosphate cleavage) with tangible biochemical consequences (specifically, which areas are most likely to intercept signals from hydrolysis and impair the effector-recruiting interfaces characterizing the GTP-active state).

To recapitulate, the above descriptions of each method and its nuances should provide the reader with an initial impression of the additional advantages that their joint use could introduce. We better contextualize such benefits in the subsections that follow, wherein we dissect trends in the normalized scores shown in Figure 6, relating them, entirely *a posteriori*, to the most recent experimental and computational allosteric understanding of K-Ras4B in its active state.[38,42,45,67] This represents the most significant final validation emerging from the integration of our methods. Where possible, we will stress which areas of the protein show universal consensus across languages, and which ones do not and why.

Allosterically Compact G Domain. Large-scale mutagenesis experiments by Weng et al.[42,45] support an allosteric model characterized by the compactness of the G domain and allosteric coupling across central β-sheets.

Our normalized scores in Figure 6 appear to fully corroborate this observation. Indeed, consensus across all scores is invariably observed at the cusp of the $\alpha3-\beta5$ loop and $\beta5$ sheet or, otherwise put, at the hub centered on residues 110−112 also identified in the preceding subsections. At equilibrium, we observe high allosteric coupling (*i.e.*, SPM prominence; light) and very low fluctuations (*i.e.*, DF; dark); in addition, there clearly emerges a low susceptibility to perturbation (*i.e.*, D-NEMD, simulations ATD; dark). Comparable consensus patterns are observed at the *N*-terminus of the $\beta4$ sheet and at the central and *C*-terminal parts of helix $\alpha5$. Interestingly, we also observe consistencies for the remaining central $\beta6$ sheet: in this case, however, scores in the "purple" range (Figure 6) identify it as an allosterically "intermediate" area, that both retains some degree of allosteric relevance at equilibrium but is also not as immune to allosteric impulses from GTP hydrolysis or heating as the other parts of the G domain mentioned above.

In light of the above, our combined "linguistic" evidence does more than simply confirm G domain compactness detected by experiment:[42,45] it shows that different areas of the G domain contribute differently to this compactness and allosteric signaling.

Recognizing Known Mutations. Importantly, the most structurally disruptive allosteric mutation sites emerging from the work of Weng et al.[42] (*i.e.*, eight "novel" allosteric nodes not responsible for GTP binding and scattered across the G domain and switches) can be directly cross-compared against our SPM. A similar comparison with the SPM can be drawn with the most abundant clinically relevant cancerous mutations of K-Ras4B.[43] The two categories expectedly show a marginal overlap, since disruptive mutations tend to destroy the GTP-active state, while cancerous ones tend to prolong its lifetime. We perform such comparison in Figure 7, wherein the SPM (Figure 3) is redrawn in off-white, except for sites corresponding to disruptive mutations[42] (shown in ruby red or pink; *vide infra*), cancerous mutations[43] (lime green), or both (purple).[42,43]

In this aspect too, focusing only on SPM data for now, our findings fully agree with (and are thus validated by) experiment. Six out of eight allosterically disruptive mutation sites[42] are

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

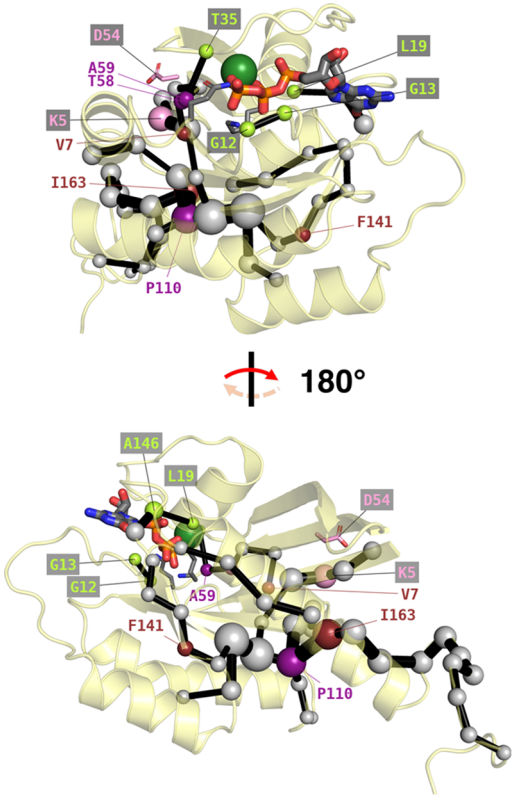


Figure 7. Alternative view of the SPM (Figure 3) from identical angles, but with 7 out of 8 "novel" allosteric residues uncovered experimentally (the 8th, Gly10, is omitted).[42] The six that directly appear on the SPM are either labeled in purple, if they also correspond to known cancerous mutations,[43] or in ruby red otherwise. Asp54, which does not lie on the SPM but interacts with Lys5 (pink), is rendered as sticks and itself labeled in pink. In lime green, we label other SPM residues closer to the active site, which frequently mutate in cancers.[43] All other SPM residues are denoted by off-white spheres.

found to be directly located on the SPM (purple, ruby red, pink in Figure 7). A seventh site,[42] Asp54 on β3, while absent from the SPM, interacts electrostatically with SPM node Lys5 (β1) (and is thus shown in pink in Figure 7). The list of residues and their locations is reported in detail as Supporting Information in the section devoted to the SPM.

Moreover, Figure 7 shows that the SPM fares equally well in mapping key clinical oncogenic mutation sites.[43] We have already discussed Pro110 in ruby red, since it is also disruptive and within the allosteric hub; Thr58 and Ala59, also in ruby red, are also disruptive and as part of switch II will be discussed later. Among remaining oncogenic mutations, in lime green, we should first and foremost mention the appearance of Gly12 and Gly13 on the P-loop, which are by far the most frequently mutating residues in oncogenic K-Ras4B.[38,43] However, we also note the presence of less frequently mutating residues,[43] namely, Leu19 on helix α1 (for guanosine binding), Thr35 on switch I (discussed at length later), and Ala146 on the β6−α5 loop (within a E*x*SAK motif notoriously conserved[40,57] in the RAS family). All such "lime green" mutation sites belong to sections of the SPM that have a moderate "shortness" (lower allosteric prominence), as opposed to the very high "shortness" characterizing the disruptive mutation sites.

Here, we also briefly focus on the P- and β6−α5 loops, and helix α1 and the definition of their allosteric roles not only in terms of SPM but also in terms of the other three "languages"; switches I and II will be dissected later. The S_{norm} heatmap in Figure 6 shows that, for the P-loop, DF and D-NEMD simulations match the SPM in painting a profile of "moderate allosteric relevance" that is similar to the one for the β6 sheet discussed previously. At equilibrium (DF matrix; Figure 2), there is clearly allosteric coordination with the G domain, as signaled by low (white) scores, and in the DF video in the Supporting Information (residues 10−14). However, it is striking to note the very low ATD S_{norm} scores (black band in Figure 6 bottom row; blue in Figure 5), suggesting little P-loop−GTP dialogue prior to hydrolysis: only *after* hydrolysis, moderate S_{norm} scores for the D-NEMD simulations suggest that there is coupling between the P-loop and the cleaving P_i. Similarly, for the Leu19 region of helix α1 and the Ala146 region of the β6−α5 loop, we have a situation of relatively high allosteric communication at equilibrium *per* the DF scores (Figures 2 and 6) that matches moderate SPM prominence, and both the ATD and D-NEMD approaches agree that there is a fair degree of allosteric signaling upon GTP perturbation.

In summary, our combined "languages" again provide us with a nuanced characterization of the dynamic "fingerprint" of different kinds of mutation sites that would have been difficult to obtain otherwise. Indeed, as oncogenic mutations are expected not to obliterate allosteric communication pathways like the other (disruptive) mutations in Figure 7 (Pro110, which suffers frameshift mutations,[43] is a more complicated case), we consider it encouraging that regions outside the switches containing cancerous mutations exhibit a clear but moderate allosteric prominence at equilibrium. If one includes the extra evidence from the D-NEMD simulations (*i.e.*, that there *is* some perturbation upon hydrolysis, particularly in Leu19 and Ala146 that are on the guanosine end of the GTP; Figure 6), this further suggests that oncogenic mutations could enhance K-Ras4B activity by mitigating the allosteric effects linked to its inactivation (GTP hydrolysis); this is certainly proven for inhibitor-resistant mutations of Ala146.[57] In the case of the P-loop, which similarly only dialogues with the G domain at equilibrium but is in contact with the reactive triphosphate moiety of GTP and is only perturbed by it *after* hydrolysis, one could envisage this region as one of the allosteric hubs that controls reactivity. It would normally receive instructions from the G domain to initiate GTP hydrolysis, but mutations of Gly12 and Gly13 could disrupt these instructions and/or mitigate the allosteric effects of hydrolysis. In the case of known mutations to Asp12 and Asp13, as discussed later, effects on hydrolysis could even be purely electronic in nature, with a much more "direct" influence on reactivity.

Finally, before we focus on mutation sites on the switches, we once again bring the reader's attention toward the stark contrast of these "blander" cancerous profiles with the much "starker" fingerprints of the allosterically disruptive mutations occurring on the central allosteric hub.

Allostery and Mutations of Switches I and II. On top of harboring relatively abundant cancerous mutation sites[43] such as Thr35, Thr58, Ala59 (Figures 3 and 7), and Gln61, it is well-established that switches I and II represent the key trait distinguishing the GTP-active state of K-Ras4B from the GDP-inactive one (Figure 1).[37,38,40] Moreover, even in the GTP-bound state, switches are reported[37,38,40,45,54] to assume several conformational states that ultimately enable recruitment of regulatory proteins such as RAF1 and GAP.

It is now a good point to focus our discussion on how our simulations identify these crucial switch regions. In this case too,

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

we find that all four methods provide very eloquent findings: indeed, the majority of switches I and II shows very low allosteric coupling to the G domain core at equilibrium (Figure 6; dark SPM; light DF) and are clearly identified by both ATD and D-NEMD simulations as allosteric hotspots for perturbative events (lightest shades, bottom two rows). Significantly, however, there is one substructure on each switch that ostensibly bucks these trends, showing up as distinctly colored bands in Figure 6 for all four methods: on switch I, this is the central portion that comprises the clinically[43] (and catalytically)[41] crucial Thr35 (Figure 7), whereas, on switch II, it is the *N*- terminus that harbors the clinically[43] (and catalytically)[41] crucial Thr58, Ala59 (Figure 7), and Gln61. These residues retain a modest degree of allosteric coupling to the allosteric hub: all have comparatively lower DF scores, and all except Gln61 feature in the SPM.

Consistency of these findings with experimental evidence is patent. The substantial flexibility of most of switches I and II is entirely in line with that reported previously.[37,38,40,45,54] Still, while switches I and II remain expectedly very mobile in the GTP-active state, we find that they both retain mild allosteric coupling to the G domain in catalytically crucial regions (cf. Figure 2 bottom): Mg^{2+}-bound Thr35 in the former case, and Thr58/Ala59/[Gln61] in the latter case. These characteristics also confer the switches control on GTP hydrolysis itself. In turn, both switches are recognized to be clearly perturbed by hydrolysis: this unequivocally suggests that once hydrolysis occurs, and departure of P_i and Mg^{2+} severs all allosteric links, switches are set to lose their ability to attain active-state conformations (eventually reaching the status in Figure 1c), and their affinity for effectors will be compromised.

In other words, our simulations confirm that there is a clear connection between GTP hydrolysis and switch deactivation, while prior to hydrolysis, crucial parts of each switch ensure an allosteric connection with the G domain. Establishing this sort of chronological distinction in a system is again only made possible by the combination of more than one method.

Allosteric Control on GTP Hydrolysis. We have seen that catalytically relevant residues Thr35, Thr58, Ala59, and Gln61, which are also subject to pathogenic mutations,[43] are located on special subportions of switches I and II that all our chosen languages pinpoint. Similarly, we have discussed our consistent findings for mutation-prone Gly12 and Gly13, which are located in the P-loop in the vicinity of the triphosphate moiety of GTP. To these residues, we should add the final catalytic actor, Lys16, which is just downstream of the P-loop, and which we have explicitly marked from Figure 2 to Figure 6. A reference for the hydrolytic mechanism, including on the role played by GAP and its Arg789 (Figure S1b),[70] is the EVB investigation by Calixto et al.[41] In this view, prior to hydrolysis, Gln61 helps position the nucleophilic water to attack GTP:Pγ, and both prior to and during hydrolysis, cationic Lys16 lowers the free energy barrier by abstracting negative charge from the transition state.

Strikingly, in our simulations, the only allosterically relevant aspect of Lys16 is general allosteric coupling with the G domain detected in the DF matrix (Figure 2 bottom, Supplementary Video). This suggests that while crucial for catalysis, Lys16 did not evolve to become as prominent an allosteric regulator of it.[41] With Gly12, Gly13, Thr35, Thr58, and Ala59 featuring in the SPM (Figure 7) but at the same time Lys16 and Gln61 excluded from it, and with DF scores showing a moderate degree of allosteric coupling to GTP at equilibrium, it is difficult to univocally establish exactly how each mutation will hamper GTP hydrolysis, or indeed, in unmutated K-Ras4B, which of these residues will mostly control hydrolysis through electronic/electrostatic intervention as opposed to acting through subtle allosteric regulation of the active site.

Of course, it is not implausible that known[43] mutations of Gly12, Gly13, or Gln61 to Asp would be more likely to intervene electronically, as in all three cases the Asp would be within reach of sequestering Arg789 in GAP (Figure S1b)[70] from GTP:Pγ. Similarly, mutation of Thr35 to Ala[43] could impede Mg^{2+} binding, thus leaving GTP unchelated and unable to hydrolyze. As also recognized by the EVB study,[41] Gly13 and Ala59 too can exert an electrostatic influence on the reaction barrier and therefore on K-Ras4B reactivity. On the other hand, vicinal residues such as Gly12 and Thr58 could indeed be controlling reactivity through allosteric effects on the active site alone. Proving these hypotheses, however, would require additional simulations and experiments and is beyond the scope of this work.

Allosteric Control of Other Functional Interfaces. Since switches I and II are the main effector recruiters in GTP-active K-Ras4B, it is encouraging to see that D-NEMD simulations (Figures 4 and 6) identify both switches to be the two areas in K-Ras4B that are most perturbed by GTP hydrolysis. However, we should assess performance of our methods on other interfaces formed by K-Ras4B, which of course extend beyond the switches. Conveniently, a number of these have been structurally characterized, including the one with the effector RAF1 (PDB ID: 6xi7; Figure S1c),[34] and the one with hydrolysis-triggering GAP (Figure S1b; PDB ID: 1wq1).[70] In their mutagenesis study, Weng et al.[42] further dissect five other K-Ras4B interfaces in full.

If we consider the above experimental evidence,[34,42,70] we observe that the agreement between D-NEMD simulations and key structurally known interfaces formed by K-Ras4B is much more far-reaching. Taking the RAF1 interface as an example (Figure S1c),[34] we see that Asp153 on the *N*-terminus of helix α5 forms an integral part of that interface; in full agreement with this, the *N*-terminus of the otherwise unperturbed helix α5 (*vide supra*) is the only part of the helix for which D-NEMD simulations show some degree of perturbations. Further still, in all remaining areas of that interface outside switch I, D-NEMD simulations (Figures 4 and 6) also indicate large perturbations *post*hydrolysis, again showing that the event impairs all parts of the interface and is consistent with obliterated RAF1 binding by the GDP-inactive state.

The ability of the D-NEMD approach to capture interface perturbations in general[17] is further confirmed by our own simulations insofar as it additionally shows perturbations along helix α1, on loop α1−β2 aside from switch I, sheet β2, and the β2−β3 loop with the initial portion of sheet β3. Indeed, in the interfaces dissected by Weng et al.,[42] the six residues that are common to all and most impacted by their mutation (Gln25, Asp33, Ile36, Asp38-Tyr40; marked in Figure S1c) are precisely located in helix α1, loop α1−β2/switch I, and sheet β2, signaling that the perturbation of this wider area during hydrolysis (and not just switch I) has actual biological repercussions. In addition, we recall that an inactive K-Ras4B (Figure 1c) also entails open β2 and β3 sheets,[38] and that even the interface with GAP (Figure S1b)[70]—which requires a certain preference for the GTP-active state even if GAP is not an effector—spans those same areas that D-NEMD simulations show as perturbed. In addition, GAP also interfaces (Figures 4, 6, and S1b) with the equally perturbed switch II and loop β3−α2, and other more marginally perturbed but crucial areas including, notably, the P-loop.

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

$\alpha3-\beta5$ Loop as a Secondary Allosteric Switch. We have amply seen that Pro110 ($\alpha3-\beta5$), Met111 ($\beta5$), Val112 ($\beta5$), and Ile163 ($\alpha5$) are of cardinal allosteric importance for maintaining the structure of the G domain (Figure 3); unsurprisingly, the importance of loop $\alpha3-\beta5$ and helix $\alpha5$ is highlighted in a number of experimental contexts,[34,40,42] a number of which—particularly for helix $\alpha5$—we have already discussed. Indeed, on top of harboring allosteric hub residue Pro110 at its *C*-terminus, the *N*-terminus of the $\alpha3-\beta5$ loop is a secondary allosteric switch that regulates the "kink" in helix $\alpha3$ (cf., *e.g.*, Figure 2), which in turns helps keep switch II in the active state alongside GTP.[40]

In this respect, it is very interesting to notice that as part of the allosteric hub, the S_{norm} profile of the *C*-terminal portion of the $\alpha3-\beta5$ loop resembles the $\beta5$ sheet itself (Figure 6); on the other hand, the immediately preceding *N*-terminal portion of the loop, along with helix $\alpha3$ itself, exhibits an opposite profile (Figure 6), with low allosteric communication at equilibrium (lighter DF, darker SPM), but clearly more sensitivity to hydrolysis (lighter D-NEMD). This is exactly the profile one would expect for a secondary switch that is triggered by GTP hydrolysis and shows once again the potential level of insight that our combined methods could provide for a system whose allosteric properties are less understood—without necessarily taking different conformational states into account.

HVR——Flexible but Fundamental. We know, from experiments,[37] of the dialogue occurring between the HVR and the G domain, while HVR flexibility is testified by its irresolution in most crystal structures.[38]

First of all, we should note in passing that while our own MD simulations do not replicate previously reported[46,47] HVR–switch (and G domain–membrane) contacts, they still reproduce mobility of the HVR across the surface of the lipid bilayer, which is in line with the established role[37] of K-Ras4B in recruiting effectors to the cellular membrane and bringing them together for dimerization.

Even more importantly, Figures 2–4 show that the HVR is approached quite differently by our chosen languages, enabling us once again to extract precious additional information. DF analysis points to high flexibility/poor allosteric coordination (intense blue color in Figure 2), and is thus in line with crystallographic evidence,[38] while the SPM (Figures 3 and 7), by mapping almost the entirety of the HVR, still confirms the important allosteric role[37] that it can play despite its flexibility. Another related finding from the SPM (Figure 3) that is of utmost significance is that Ile163 on helix $\alpha5$ is the fundamental residue that intercepts allosteric signals from the HVR and relays them into the G domain. Further still, after hydrolysis, D-NEMD simulations detect very little deviation in the HVR: on top of being a testament to the better ability of D-NEMD to cancel out nonallosteric noise, this intriguingly suggests that allosteric perturbations from hydrolysis are not wasted by dissipating back into the membrane.

In this case too, in addition to reproducing the fundamental characteristics of HVR, our joint "multilingual" approach has demonstrated the capability of capturing different aspects of it (flexibility vs allosteric communication) that would have been otherwise captured only partially. In addition, it has crucially enabled us to discover the importance of Ile163 in relation to the HVR, and how the latter is unaffected by hydrolysis.

Allosteric Pockets. We finally extend our analysis to the four allosteric pockets I–IV previously recognized by Grant et al. through simulations[44] and later validated experimentally,[42] to evaluate how well our recipe can recognize these pockets and thus assess its potential in predicting allosteric pockets in new systems. (In the designation used by both studies,[42,44] inhibitors Sotorasib and BI-2865 both bind to pocket II). While interference with all four pockets ultimately ushers in perturbed K-Ras4B activity, our own assessment reveals that pockets II,[44] IV,[44] and I/III[44] fall into three distinct categories of which the latter is distinct from the former two.

More specifically, pocket II should be envisaged as a stabilizer of the GDP-inactive state (Figure 1c),[44] which occupies the volume created between inactivated switch II and allosteric helix $\alpha3$. The shallow pocket IV is found[44] through blind docking on GTP-active K-Ras4B and occupies the volume between switch I, helix $\alpha1$, and the rest of sheet $\beta2$. These regions correspond to the interface with RAF1 discussed earlier (Figure S1c) and indeed, the purpose of targeting this pocket would be to disrupt the K-Ras4B–RAF1 interface. In this case, since disruption of the interface is captured by the D-NEMD approach, one could assume that application of our approaches to other systems could automatically help identify interfaces that could be therapeutically targeted. However, we must point out that identification in a system of pockets such as pocket IV would require some prior degree of knowledge about interfaces to understand which ones are therapeutically targetable and which are not. Similarly, identifying pockets like pocket II would require structural knowledge of other biologically relevant conformational states, and explicitly taking them into account.

On the other hand, pockets such as I/III specifically disrupt crucial allosteric sites within the GTP-active state. This means that they are fully identifiable by our chosen approach: as clearly inferable especially from SPM (Figure 3) and D-NEMD data (Figure 4), both these pockets occupy regions of utmost allosteric importance at equilibrium in the active state, while feeling little to no perturbation from GTP hydrolysis. More specifically, pocket I is located in the coupled region between sheets $\beta1$, $\beta3$, and the *C*-terminus of helix $\alpha5$; and pocket III spans none other than the main $\alpha3-\beta5/\beta5/\alpha5$ allosteric hub. Interference with these pockets is thus bound to bring maximum disruption to the active state, and experimental data certainly suggests so too.[42]

This correspondence with experiment bodes well for novel systems whose allosteric properties are not as pervasively known, and wherein joint application of our chosen approaches could be synergistically beneficial in driving the design of novel allosteric modulators. Indeed, even when used on their own, DF analysis,[95] SPM,[28] and the D-NEMD approach[67] have already shown significant promise in the (direct or indirect) identification of allosteric sites. Use of ATD for similar purposes has been equally endorsed by its authors.[68] Concretely, one could, for example, compare and contrast the information provided by the various languages and predict likely allosteric sites as clusters of adjacent residues for which there is a degree of consensus: this would likely disrupt allostery if interfered with by a ligand.

Final Picture: The Membrane Talks to GTP, GTP Talks to the Switches. In summary, after piecing together information from our own simulations and from previous work, there emerges a chronologically clear, articulate, and experimentally consistent allosteric portrait of K-Ras4B. For the chosen GTP-active state of K-Ras4B[34] (*vide infra*), languages consistently describe a compact protein in which allosteric signals appear to be far-reaching but rigorously compartmentalized: prior to hydrolysis, at equilibrium, a set of pathways

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901–919

affords a more or less strict control on hydrolysis itself while retaining a mild dialogue with the switches; once hydrolysis begins, on the other hand, signals mainly (but not exclusively) propagate to effector switches I and II, disrupting their function. Such a narrative would have clearly been impossible to reconstruct if the languages had been used separately.

More specifically, at equilibrium, helix α5 and sheets β1, β4−β6, as a rigid core, constitute the allosteric centerpiece of the G domain; in particular, the high allosteric significance of sheets β4 and β5 is captured by all four allosteric languages, and that of helix α5 by three of them out of four. Despite its intrinsic flexibility, it is also coherently found that the HVR is paramount in relaying signals between the membrane and the above rigid core. While the intricate network of allosteric dialogue ostensibly avoids the majority (but crucially not the entirety) of switches I and II, there are convincing signals from SPM, and to a certain extent DF and ATD, that there does exist a direct allosteric coupling going from the HVR all the way up to key binding site residues, including catalytic ones. The said binding site residues are first and foremost the P-loop and guanine-binding residue Ala146 on the β6−α5 loop (G5 motif).

Albeit more blandly, however, parts of both switches are also involved in the allosteric control: switch II through Thr58/Ala59 (DF and SPM) and switch I through Thr35 (DF and SPM). Existence of these allosteric links at equilibrium to both switches, however mild, is the fundamental reason why switches can mediate interaction with regulatory proteins only in the active state. Similarly, the electronic relevance[41] of some allosterically important active site residues not within switches suggests that allosteric control on GTP hydrolysis is likely to be exerted for the greater part through these residues rather than through Lys16 and Gln61, which influence the reaction more "directly", through electronic effects.

Moving on along our timeline of events, anyhow, once hydrolysis is initiated, the strong allosteric link between the cleaving phosphate center and switches I and II (with Gln61)—which get visibly perturbed (D-NEMD; Figure 4) and of course mediate K-Ras4B inactivation—could not be clearer. There also is an unequivocal disruption of the secondary switch on the *N*-terminus of the α3−β5 loop; of other regions forming interfaces, including the *C*-terminus of helix α1, the β2 sheet outside switch I, and loops α1−β2 and β2−β3; and, again, of the β6−α5 loop.

While in proximity to the hydrolytic center, P-loop and Lys16 remain comparatively less perturbed by hydrolysis. It is also important to recall that the central β-sheet core, α5 helix, and HVR are entirely decoupled from hydrolysis, even despite the latter's flexibility, meaning signals from hydrolysis are unlikely to be relayed back to the membrane: this serendipitous finding actually makes biological sense, insofar as the GTPase should have evolved in such a way that little or none of the energy resulting from hydrolysis is (wastefully) dissipated away from biologically functional areas.

Not just Four (Allosteric) Languages. Having thus far ascertained the encouraging agreement between our four "languages" and experimental data, a crucial point to consider are the reasons and implications of our methodological choices. In the much broader context of allosteric studies, we should obviously start by stressing that plenty of other options would have been available for our investigations. While it is not possible to review them here in detail, the reader can certainly find some excellent accounts elsewhere,[5,21,23,25] including in a special 2022 issue of *J. Mol. Biol.*[25] dedicated to allostery. Disregarding for a moment whether or not the reference MD simulations are atomistic and/or unbiased, for instance, the (highly recommendable) editorial review presenting the issue[25] provides a very informative overview of the individual studies published therein, which include a wide range of allostery detection methods. These include various flavors of Gaussian Network Models,[96] DF analysis,[16] and a range of other sophisticated approaches based on machine learning and coevolutionary approaches. In consulting these approaches, one can easily recognize similarities, differences, and elements of uniqueness with respect to DF, SPM, D-NEMD simulations, and ATD.

Additional examples taken from across the realms of equilibrium and perturbative allostery alike include Normal Mode Analysis,[97] Leverage Coupling,[21] Perturbed Ensemble Analysis,[98] and Networks of Local Correlated Motions.[99] To these, one must add online resources.[5,21,23,25] It is also important to reiterate that findings from a chosen combination of languages could be further reinforced by forms of systematic coevolution analysis[2,27] (such as those mentioned in ref 25) and/or analysis of the most frequent pathogenic mutation sites.[43] Any sufficiently variegated combination of these options should, by all means, be considered as alternatives by any reader wishing to plan their own investigation.

Given this elevated number of valid "allosteric languages" available across the scientific community, it was clearly beyond the scope of this work to exhaustively compare and contrast as many of them as possible. Rather, we have opted to limit ourselves to four methods—quite different in scope and theoretical origins—and not "compare" them, but use them as extensively as possible on a well-documented oncotarget, with the aim of highlighting just how many subtleties there are to allostery, and how powerful a suitably planned combination of (any) *in silico* allostery detection methods can be in capturing them adequately.

Before putting our own choice of methods into perspective (cf. next subsection), we should, in any case, stress that our choice of methods does not imply an indication of preference or support for any one of them: in fact, we have mainly aimed to seek methods that were as heterogeneous as possible and beyond our immediate areas of expertise. Much like learning new languages costs time but is worth the investment, we recommend to authors wishing to undertake a new allosteric study that they too seek to maximize heterogeneity in their chosen methods, while monitoring key aspects such as noise, portability, computational cost, and user-friendliness. As Supporting Information, we present some of the considerations specifically applying to our own choice of methods and review them critically.

Not just One Target. In light of the experimental validation made earlier, we can make a solid argument that when applied and interpreted rationally, the four allostery detection methods featured in this study—all of which are automatable—can work synergistically to provide a high level of allosteric detail about that they would be unable to provide separately. Furthermore, despite having here only worked with K-Ras4B as a benchmark target, we are in fact confident about the portability of our approach to other targets too; this is, of course, a crucial aspect to address, given the sheer heterogeneity of targets and contexts in which allostery could be of interest.[25] We are first and foremost convinced of the portability of our approach in light of the fact that DF,[14,95] SPM,[12,28] D-NEMD simulations,[17,63−67] and ATD[68] alike are known to have fared very well when individually applied to a variety of targets, some of which significantly larger than K-Ras4B.

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

Always within the context of portability, a further fundamental advantage of our computational recipe that is worth reiterating is its total applicability to unbiased MD simulations, without the need to sample large-scale conformational changes or start from alternative conformations: after all, the large amount of data originating from this work is, essentially, derived from just a single starting structure. Further to this point, and returning to the general context of allosteric studies, though our very heterogeneous choice of methods only represents a fraction of those available (*vide supra*), it is indicative that we still were able to make all languages work so articulately: simply combining areas of consensus vs disagreement into a sort of "metalanguage" was enough to bring to light some strikingly consistent allosteric traits. Inexistence of initial biases or hypotheses is another clue to the portability of this approach to other systems, which only becomes limited by computational resources. Alternative combinations of methods to study different targets should, again, be chosen so that they retain the same advantages.

SUMMARY AND CONCLUSIONS

Allostery has evolved alongside proteins to regulate most aspects of their biological function, enabling residues that give rise to functional interfaces, pockets and/or active sites to dynamically influence each other even when they lie tens of Ångströms apart. Perturbations of the delicate allosteric equilibria governing proteins can have far-reaching and oftentimes detrimental effects, for example, ushering in pathogenic alterations of reactivity and promotion of harmful interactions over beneficial ones. The multifaceted ways in which allostery can manifest itself during the lifecycle of a protein resemble spoken "languages": speaking them correctly (*i.e.*, modeling them accurately) should be of great help in understanding molecular mechanisms driving (changes in) biological function of a particular protein. Unbiased molecular dynamics (MD) simulations are an option but require decryption with a suitable allostery detection method (language) due to the sometimes too long operational time scales of allostery; while several such (very valid) methods exist, they tend to be used in isolation by their reference community, thus affording a partial, "monolinguistic" picture.

In this work, we have conducted a series of 20 independent unbiased MD simulations of the fully solvated, membrane-embedded, unmutated GTPase K-Ras4B, in an "active" GTP-bound state that is characterized by two allosteric switches poised to recruit effector proteins. Aiming to prove our argument that the combination of more than one allosteric language should appreciably improve the allosteric model of a system under study, we proceed to decrypt allostery in these simulations using four representative allostery detection methods (languages). The intensely studied oncoprotein K-Ras4B,[37,38,42] with its notoriously difficult pharmaceutical targetability, was deliberately chosen to prove the benefits of this combination of methods.

Languages chosen to interrogate the active state of K-Ras4B at equilibrium include: (i) distance fluctuation (DF) analysis,[14,16,61] which postulates that pairs of residues moving more in a concerted manner than others—with distances remaining closer to the simulation average—should represent hotspots of allosteric change; (ii) the shortest path map (SPM),[12,31] which reconstructs the main allosteric communication pathway based on networks of vicinal residues that move with high (anti)-correlation. Out of equilibrium, (iii) dynamical nonequilibrium MD (D-NEMD) simulations,[17,62−67] which spawn a large number of short (50 ps) MD simulations from as many unperturbed MD frames, in which GTP is nearly instantaneously hydrolyzed by brute force: deviation from equilibrium MD, averaged over all short simulations, reflects the degree of allosteric perturbation induced by hydrolysis. Finally, (iv) anisotropic thermal diffusion[68] supercools a series of frames isolated at regular intervals, equilibrates them, and then reheats the GTP only: deviation with respect to the first frame reflects the degree of coupling to the active site.

As expected, the four methods paint a very articulate allosteric picture of GTP-active K-Ras4B. Owing to their ability to capture different expressions of allostery, while all four languages concur in correctly identifying high or low allosteric importance for certain areas of K-Ras4B, they provide intriguingly different answers for certain other areas: thanks to these linguistic nuances one can provide a clear chronological dimension to K-Ras4B allostery. At equilibrium, prominent allosteric communication pathways travel from the membrane and through the flexible hypervariable region, from when they reach a paramount allosteric hub centered around the *C*-terminal half of (tendentially rigid) helix $\alpha5$, sheet $\beta5$, and around loop $\alpha3-\beta5$. From here, they branch out to encompass most remaining parts of the protein, notably including—albeit to moderate degrees—residues on several sides of the GTP binding site. Among these are residues on both switches (Thr35 on I and Thr58/Ala59 on II) and mutation-prone Gly12 and Gly13 on the P-loop. As shown by EVB reactivity studies,[41] these residues or others in their immediate vicinity, such as Lys16 and Gln61, exert some form of control on GTP hydrolysis, signifying that the active state can withhold its own inactivation regardless of the presence of the GTPase stimulator GAP. Despite the above allosteric links, switches exhibit low allosteric coordination, in agreement with their interfacial plasticity. Upon hydrolysis, most of K-Ras4B remains unaffected, except for the two switches (on which it is known to have an inactivating impact), and the $\alpha1/\alpha1-\beta2/\beta2$ area: all belong to crystallographically known interfaces.[34,70]

Consistency with experiment is recognizable in a number of other aspects, notably in the importance of the $\alpha5/\beta5/\alpha3-\beta5$ allosteric hub,[42] general allosteric compactness,[42,45] and retention of allosteric links in the hypervariable region and in both switches.[38] Our data also agree with experiment in terms of the allosteric relevance of helix $\alpha5$ and the G5 domain;[57] and two of four pockets suggested by mutagenesis,[42] which, rather than stabilizing the GDP-inactive state, are likely to disrupt the GTP-active state instead, are also located in allosterically active regions.

To conclude, while notably different in conception, pitfalls, and genesis (exactly like spoken languages), the four chosen allosteric languages in synergy have depicted a very nuanced and experimentally consistent allosteric portrait of K-Ras4B that they would have been unable to provide if applied on their own. Crucially, decryption of such an articulate portrait was possible even if our simulations all began from a single structure in its GTP-active state. The coherent results produced by our chosen techniques on allosteric pathways in K-Ras4B provide elegant reconfirmation of allostery as a universal property of proteins, from therapeutic targets for allosteric modulators to biocatalysts, regardless of the computational "languages" used to decipher it. We believe our work proves the benefits of applying as many "allosteric languages" as computational resources permit.

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.3c11396.

Depiction of the starting structure and the K-Ras4B—GAP complex; MD preproduction details; alignment, clustering, and matrix derivation for SPM; full SPM branch details; allosteric mutation sites appearing in the SPM; further D-NEMD details and "reactive" pose counts; derivation of S_{norm} scores; critical assessment of the chosen methods (PDF)

D-NEMD: average Cα deviation from 0 to 50 ps, projected onto the starting structure (MP4)

DF analysis: intensity of DF scores for all possible K-Ras4B residue pairs, highlighted on the starting structure, progressing from residue 2 to residue 185 (MP4)

Equilibrium MD: starting coordinates, topology, input; ATD: input; D-NEMD: topology and input; SPM: alignment scripts, input matrices, output (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Stefano A. Serapian − *Department of Chemistry, University of Pavia, 27100 Pavia, Italy;* orcid.org/0000-0003-0122-8499; Email: stefanoartin.serapian@unipv.it

Giorgio Colombo − *Department of Chemistry, University of Pavia, 27100 Pavia, Italy;* orcid.org/0000-0002-1318-668X; Email: g.colombo@unipv.it

Authors

Matteo Castelli − *Department of Chemistry, University of Pavia, 27100 Pavia, Italy;* Present Address: Department of Chemistry and Biochemistry, University of California, San Diego, 4218 Urey Hall, 9500 Gilman Dr., La Jolla, California 92093-0340, United States

Filippo Marchetti − *Department of Chemistry, University of Pavia, 27100 Pavia, Italy; INSTM, 50121 Florence, Italy; E4 Computer Engineering, 42019 Scandiano (RE), Italy*

Sílvia Osuna − *Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Girona, Catalonia E-17071, Spain; ICREA, Barcelona, Catalonia E-08010, Spain;* orcid.org/0000-0003-3657-6469

A. Sofia F. Oliveira − *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.;* orcid.org/0000-0001-8753-4950

Adrian J. Mulholland − *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.;* orcid.org/0000-0003-1015-4567

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.3c11396

Funding

M.C., F.M., S.A.S., and G.C. thank AIRC (IG 27139), PRIN (Grant 20209KYCH9), and Programma di ricerca CN00000013 "National Centre for HPC, Big Data and Quantum Computing" for funding; they also acknowledge "Funzione Salute ENI" for generous allocation of computing resources. S.O. thanks the Generalitat de Catalunya for the consolidated group TCBioSys (SGR 2021 00487) and Grant Projects PID2021-129034NB-I00 and PDC2022-133950-I00 funded by Spanish MICIN. S.O. is also grateful to the funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-2015-StG-679001, ERC-2022-POC-101112805, and ERC-2022-CoG-101088032) and the Human Frontier Science Program (HFSP) for Project Grant RGP0054/2020. This work is part of a project that has received funding from the European Research Council under the European Horizon 2020 research and innovation program (PREDACTED Advanced Grant Agreement No. 101021207) to A.J.M. The authors A.S.F.O. and A.J.M. thank the Biotechnology and Biological Sciences Research Council (BBSRC Grant Numbers BB/W003449/1, BB/L01386X/1, and BB/X009831/1); A.S.F.O. also thanks the BBSRC for her Discovery Fellowship, and Oracle for Research for her research fellowship.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

S.A.S. thanks Prof. Alemayehu Gorfe (University of Texas Medical School—Houston) for providing an initial structure of membrane-bound K-Ras4B, Prof. Ana Vila Verde (University of Duisburg-Essen) for assistance with her inorganic phosphate force field parameters, and Prof. Aleksander Lyubartsev (Stockholm University) for assistance with the *Slipids* force field. The authors at the University of Pavia acknowledge support from the Italian *Ministero dell'Università e della Ricerca* (MUR) and the University itself through the program *Dipartimenti di Eccellenza 2023−2027*.

■ REFERENCES

(1) Lockless, S. W.; Ranganathan, R. Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science* **1999**, *286* (5438), 295−299.

(2) Granata, D.; Ponzoni, L.; Micheletti, C.; Carnevale, V. Patterns of coevolving amino acids unveil structural and dynamical domains. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (50), E10612−E10621.

(3) Orozco, M. A theoretical view of protein dynamics. *Chem. Soc. Rev.* **2014**, *43* (14), 5051−5066. 10.1039/C3CS60474H

(4) Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R. The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem. Rev.* **2016**, *116* (11), 6391−6423.

(5) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450* (7172), 964−972.

(6) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508* (7496), 331−339.

(7) Gunasekaran, K.; Ma, B.; Nussinov, R. Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct., Funct., Bioinf.* **2004**, *57* (3), 433−443.

(8) Goodey, N. M.; Benkovic, S. J. Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **2008**, *4* (8), 474−482.

(9) Liu, J.; Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Comput. Biol.* **2016**, *12* (6), No. e1004966.

(10) Tsai, C.-J.; Nussinov, R. A Unified View of "How Allostery Works. *PLoS Comput. Biol.* **2014**, *10* (2), No. e1003394.

(11) Bunzel, H. A.; Anderson, J. L. R.; Hilvert, D.; Arcus, V. L.; van der Kamp, M. W.; Mulholland, A. J. Evolution of dynamical networks enhances catalysis in a designer enzyme. *Nat. Chem.* **2021**, *13* (10), 1017−1022.

(12) Romero-Rivera, A.; Garcia-Borràs, M.; Osuna, S. Role of Conformational Dynamics in the Evolution of Retro-Aldolase Activity. *ACS Catal.* **2017**, *7* (12), 8524−8532.

(13) Świderek, K.; Tuñón, I.; Martí, S.; Moliner, V. Protein Conformational Landscapes and Catalysis. Influence of Active Site

Conformations in the Reaction Catalyzed by L-Lactate Dehydrogenase. *ACS Catal.* **2015**, *5* (2), 1172−1185.
(14) Castelli, M.; Bhattacharya, K.; Abboud, E.; Serapian, S. A.; Picard, D.; Colombo, G. Phosphorylation of the Hsp90 Co-Chaperone Hop Changes its Conformational Dynamics and Biological Function. *J. Mol. Biol.* **2023**, *435* (3), No. 167931.
(15) Sánchez-Martin, C.; Serapian, S. A.; Colombo, G.; Rasola, A. Dynamically Shaping Chaperones. Allosteric Modulators of HSP90 Family as Regulatory Tools of Cell Metabolism in Neoplastic Progression. *Front. Oncol.* **2020**, *10*, No. 1177.
(16) Triveri, A.; Sanchez-Martin, C.; Torielli, L.; Serapian, S. A.; Marchetti, F.; D'Acerno, G.; Pirota, V.; Castelli, M.; Moroni, E.; Ferraro, M.; et al. Protein Allostery and Ligand Design: Computational Design Meets Experiments to Discover Novel Chemical Probes. *J. Mol. Biol.* **2022**, *434* (17), No. 167468.
(17) Oliveira, A. S. F.; Ciccotti, G.; Haider, S.; Mulholland, A. J. Dynamical nonequilibrium molecular dynamics reveals the structural basis for allostery and signal propagation in biomolecular systems. *Eur. Phys. J. B* **2021**, *94* (7), No. 144.
(18) Monod, J.; Wyman, J.; Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **1965**, *12* (1), 88−118.
(19) Cooper, A.; Dryden, D. T. F. Allostery without conformational change. *Eur. Biophys. J.* **1984**, *11* (2), 103−109.
(20) Cui, Q.; Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* **2008**, *17* (8), 1295−1307.
(21) Guarnera, E.; Berezovsky, I. N. Allosteric sites: remote control in regulation of protein activity. *Curr. Opin. Struct. Biol.* **2016**, *37*, 1−8.
(22) Nussinov, R.; Tsai, C.-J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153* (2), 293−305.
(23) Wei, G.; Xi, W.; Nussinov, R.; Ma, B. Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chem. Rev.* **2016**, *116* (11), 6516−6551.
(24) Gianni, S.; Jemth, P. Allostery Frustrates the Experimentalist. *J. Mol. Biol.* **2023**, *435* (4), No. 167934.
(25) Berezovsky, I. N.; Nussinov, R. Multiscale Allostery: Basic Mechanisms and Versatility in Diagnostics and Drug Design. *J. Mol. Biol.* **2022**, *434* (17), No. 167751.
(26) Nussinov, R.; Tsai, C.-J. Unraveling structural mechanisms of allosteric drug action. *Trends Pharmacol. Sci.* **2014**, *35* (5), 256−264.
(27) Crean, R. M.; Gardner, J. M.; Kamerlin, S. C. L. Harnessing Conformational Plasticity to Generate Designer Enzymes. *J. Am. Chem. Soc.* **2020**, *142* (26), 11324−11342.
(28) Maria-Solano, M. A.; Serrano-Hervás, E.; Romero-Rivera, A.; Iglesias-Fernández, J.; Osuna, S. Role of conformational dynamics in the evolution of novel enzyme function. *Chem. Commun.* **2018**, *54* (50), 6622−6634.
(29) Serapian, S. A.; Colombo, G. Designing Molecular Spanners to Throw in the Protein Networks. *Chem. − Eur. J.* **2020**, *26* (21), 4656−4670.
(30) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. Fine-Tuning Multiprotein Complexes Using Small Molecules. *ACS Chem. Biol.* **2012**, *7* (8), 1311−1320.
(31) Osuna, S. The challenge of predicting distal active site mutations in computational enzyme design. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11* (3), No. e1502.
(32) McLaughlin, S. H.; Sobott, F.; Yao, Z.-p.; Zhang, W.; Nielsen, P. R.; Grossmann, J. G.; Laue, E. D.; Robinson, C. V.; Jackson, S. E. The Co-chaperone p23 Arrests the Hsp90 ATPase Cycle to Trap Client Proteins. *J. Mol. Biol.* **2006**, *356* (3), 746−758.
(33) Bell, E. L.; Finnigan, W.; France, S. P.; Green, A. P.; Hayes, M. A.; Hepworth, L. J.; Lovelock, S. L.; Niikura, H.; Osuna, S.; Romero, E.; et al. Biocatalysis. *Nat. Rev. Methods Primers* **2021**, *1* (1), No. 46.
(34) Tran, T. H.; Chan, A. H.; Young, L. C.; Bindu, L.; Neale, C.; Messing, S.; Dharmaiah, S.; Taylor, T.; Denson, J.-P.; Esposito, D.; et al. KRAS interaction with RAF1 RAS-binding domain and cysteine-rich domain provides insights into RAS-mediated RAF activation. *Nat. Commun.* **2021**, *12* (1), No. 1176.
(35) Prakash, P.; Zhou, Y.; Liang, H.; Hancock, J. F.; Gorfe, A. A. Oncogenic K-Ras Binds to an Anionic Membrane in Two Distinct Orientations: A Molecular Dynamics Analysis. *Biophys. J.* **2016**, *110* (5), 1125−1138.
(36) Bera, A. K.; Lu, J.; Wales, T. E.; Gondi, S.; Gurbani, D.; Nelson, A.; Engen, J. R.; Westover, K. D. Structural basis of the atypical activation mechanism of KRASV14I. *J. Biol. Chem.* **2019**, *294* (38), 13964−13972.
(37) Nussinov, R.; Jang, H.; Tsai, C.-J.; Liao, T.-J.; Li, S.; Fushman, D.; Zhang, J. Intrinsic protein disorder in oncogenic KRAS signaling. *Cell. Mol. Life Sci.* **2017**, *74* (17), 3245−3261.
(38) Pantsar, T. The current understanding of KRAS protein structure and dynamics. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 189−198.
(39) Janosi, L.; Gorfe, A. A. Segregation of Negatively Charged Phospholipids by the Polycationic and Farnesylated Membrane Anchor of Kras. *Biophys. J.* **2010**, *99* (11), 3666−3674.
(40) Johnson, C. W.; Reid, D.; Parker, J. A.; Salter, S.; Knihtila, R.; Kuzmic, P.; Mattos, C. The small GTPases K-Ras, N-Ras, and H-Ras have distinct biochemical properties determined by allosteric effects. *J. Biol. Chem.* **2017**, *292* (31), 12981−12993.
(41) Calixto, A. R.; Moreira, C.; Pabis, A.; Kötting, C.; Gerwert, K.; Rudack, T.; Kamerlin, S. C. L. GTP Hydrolysis Without an Active Site Base: A Unifying Mechanism for Ras and Related GTPases. *J. Am. Chem. Soc.* **2019**, *141* (27), 10684−10701.
(42) Weng, C.; Faure, A. J.; Lehner, B. The energetic and allosteric landscape for KRAS inhibition. *bioRxiv* **2022**, DOI: 10.1101/2022.12.06.519122. (accessed 2023−12−05)
(43) Forbes, S. A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S.; et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **2015**, *43* (D1), D805−D811.
(44) Grant, B. J.; Lukman, S.; Hocker, H. J.; Sayyah, J.; Brown, J. H.; McCammon, J. A.; Gorfe, A. A. Novel Allosteric Sites on Ras for Lead Generation. *PLoS One* **2011**, *6* (10), No. e25711.
(45) Chao, F.-A.; Dharmaiah, S.; Taylor, T.; Messing, S.; Gillette, W.; Esposito, D.; Nissley, D. V.; McCormick, F.; Byrd, R. A.; Simanshu, D. K.; Cornilescu, G. Insights into the Cross Talk between Effector and Allosteric Lobes of KRAS from Methyl Conformational Dynamics. *J. Am. Chem. Soc.* **2022**, *144* (9), 4196−4205.
(46) Chavan, T. S.; Jang, H.; Khavrutskii, L.; Abraham, S. J.; Banerjee, A.; Freed, B. C.; Johannessen, L.; Tarasov, S. G.; Gaponenko, V.; Nussinov, R.; Tarasova, N. High-Affinity Interaction of the K-Ras4B Hypervariable Region with the Ras Active Site. *Biophys. J.* **2015**, *109* (12), 2602−2613.
(47) Jang, H.; Banerjee, A.; Chavan, T. S.; Lu, S.; Zhang, J.; Gaponenko, V.; Nussinov, R. The higher level of complexity of K-Ras4B activation at the membrane. *FASEB J.* **2016**, *30* (4), 1643−1655.
(48) Kapoor, A.; Travesset, A. Differential dynamics of RAS isoforms in GDP- and GTP-bound states. *Proteins: Struct., Funct., Bioinf.* **2015**, *83* (6), 1091−1106.
(49) López, C. A.; Agarwal, A.; Van, Q. N.; Stephen, A. G.; Gnanakaran, S. Unveiling the Dynamics of KRAS4b on Lipid Model Membranes. *J. Membr. Biol.* **2021**, *254* (2), 201−216.
(50) Lu, S.; Banerjee, A.; Jang, H.; Zhang, J.; Gaponenko, V.; Nussinov, R. GTP Binding and Oncogenic Mutations May Attenuate Hypervariable Region (HVR)-Catalytic Domain Interactions in Small GTPase K-Ras4B, Exposing the Effector Binding Site. *J. Biol. Chem.* **2015**, *290* (48), 28887−28900.
(51) Pantsar, T.; Rissanen, S.; Dauch, D.; Laitinen, T.; Vattulainen, I.; Poso, A. Assessment of mutation probabilities of KRAS G12 missense mutants and their long-timescale dynamics by atomistic molecular simulations and Markov state modeling. *PLoS Comput. Biol.* **2018**, *14* (9), No. e1006458.
(52) Prakash, P.; Gorfe, A. A. Lessons from computer simulations of Ras proteins in solution and in membrane. *Biochim. Biophys. Acta, Gen. Subj.* **2013**, *1830* (11), 5211−5218.
(53) Prakash, P.; Gorfe, A. A. Probing the Conformational and Energy Landscapes of KRAS Membrane Orientation. *J. Phys. Chem. B* **2019**, *123* (41), 8644−8652.

(54) Zeng, J.; Chen, J.; Xia, F.; Cui, Q.; Deng, X.; Xu, X. Identification of functional substates of KRas during GTP hydrolysis with enhanced sampling simulations. *Phys. Chem. Chem. Phys.* **2022**, *24* (13), 7653−7665.
(55) Zhang, M.; Jang, H.; Nussinov, R. The mechanism of PI3Kα activation at the atomic level. *Chem. Sci.* **2019**, *10* (12), 3671−3680.
(56) Canon, J.; Rex, K.; Saiki, A. Y.; Mohr, C.; Cooke, K.; Bagal, D.; Gaida, K.; Holt, T.; Knutson, C. G.; Koppada, N.; et al. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **2019**, *575* (7781), 217−223.
(57) Kim, D.; Herdeis, L.; Rudolph, D.; Zhao, Y.; Böttcher, J.; Vides, A.; Ayala-Santos, C. I.; Pourfarjam, Y.; Cuevas-Navarro, A.; Xue, J. Y.; et al. Pan-KRAS inhibitor disables oncogenic signalling and tumour growth. *Nature* **2023**, *619*, 160−166.
(58) Colombo, G. Computing allostery: from the understanding of biomolecular regulation and the discovery of cryptic sites to molecular design. *Curr. Opin. Struct. Biol.* **2023**, *83*, No. 102702.
(59) Dušan, P.; Kamerlin, S. C. L. Molecular modeling of conformational dynamics and its role in enzyme evolution. *Curr. Opin. Struct. Biol.* **2018**, *52*, 50−57.
(60) Verkhivker, G.; Alshahrani, M.; Gupta, G.; Xiao, S.; Tao, P. From Deep Mutational Mapping of Allosteric Protein Landscapes to Deep Learning of Allostery and Hidden Allosteric Sites: Zooming in on "Allosteric Intersection" of Biochemical and Big Data Approaches. *Int. J. Mol. Sci.* **2023**, *24* (9), 7747.
(61) Morra, G.; Potestio, R.; Micheletti, C.; Colombo, G. Corresponding Functional Dynamics across the Hsp90 Chaperone Family: Insights from a Multiscale Analysis of MD Simulations. *PLoS Comput. Biol.* **2012**, *8* (3), No. e1002433.
(62) Ciccotti, G.; Jacucci, G. Direct Computation of Dynamical Response by Molecular Dynamics: The Mobility of a Charged Lennard-Jones Particle. *Phys. Rev. Lett.* **1975**, *35* (12), 789−792.
(63) Abreu, B.; Lopes, E. F.; Oliveira, A. S. F.; Soares, C. M. F508del disturbs the dynamics of the nucleotide binding domains of CFTR before and after ATP hydrolysis. *Proteins: Struct., Funct., Bioinf.* **2020**, *88* (1), 113−126.
(64) Damas, J. M.; Oliveira, A. S. F.; Baptista, A. M.; Soares, C. M. Structural consequences of ATP hydrolysis on the ABC transporter NBD dimer: Molecular dynamics studies of HlyB. *Protein Sci.* **2011**, *20* (7), 1220−1230.
(65) Galdadas, I.; Qu, S.; Oliveira, A. S. F.; Olehnovics, E.; Mack, A. R.; Mojica, M. F.; Agarwal, P. K.; Tooke, C. L.; Gervasio, F. L.; Spencer, J.; et al. Allosteric communication in class A β-lactamases occurs via cooperative coupling of loop dynamics. *eLife* **2021**, *10*, No. e66567.
(66) Oliveira, A. S. F.; Shoemark, D. K.; Campello, H. R.; Wonnacott, S.; Gallagher, T.; Sessions, R. B.; Mulholland, A. J. Identification of the Initial Steps in Signal Transduction in the $\alpha4\beta2$ Nicotinic Receptor: Insights from Equilibrium and Nonequilibrium Simulations. *Structure* **2019**, *27* (7), 1171−1183.
(67) Chan, H. T. H.; Oliveira, A. S. F.; Schofield, C. J.; Mulholland, A. J.; Duarte, F. Dynamical Nonequilibrium Molecular Dynamics Simulations Identify Allosteric Sites and Positions Associated with Drug Resistance in the SARS-CoV-2 Main Protease. *JACS Au* **2023**, *3*, 1767−1774.
(68) Ota, N.; Agard, D. A. Intramolecular Signaling Pathways Revealed by Modeling Anisotropic Thermal Diffusion. *J. Mol. Biol.* **2005**, *351* (2), 345−354.
(69) Schrödinger, L. *The PyMOL Molecular Graphics System, Version 1.8*, 2015.
(70) Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmüller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A. The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants. *Science* **1997**, *277* (5324), 333−339.
(71) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668−1688.
(72) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525−537.
(73) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.
(74) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65* (3), 712−725.
(75) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78* (8), 1950−1958.
(76) Khoury, G. A.; Thompson, J. P.; Smadbeck, J.; Kieslich, C. A.; Floudas, C. A. Forcefield_PTM: Ab Initio Charge and AMBER Forcefield Parameters for Frequently Occurring Post-Translational Modifications. *J. Chem. Theory Comput.* **2013**, *9* (12), 5653−5674.
(77) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020−9041.
(78) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion−Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8* (4), 1493−1502.
(79) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **2003**, *24* (9), 1016−1025.
(80) Kashefolgheta, S.; Vila Verde, A. Developing force fields when experimental data is sparse: AMBER/GAFF-compatible parameters for inorganic and alkyl oxoanions. *Phys. Chem. Chem. Phys.* **2017**, *19* (31), 20593−20607. 10.1039/C7CP02557B
(81) Jämbeck, J. P. M.; Lyubartsev, A. P. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *J. Chem. Theory Comput.* **2012**, *8* (8), 2938−2948.
(82) Jämbeck, J. P. M.; Lyubartsev, A. P. Another Piece of the Membrane Puzzle: Extending Slipids Further. *J. Chem. Theory Comput.* **2013**, *9* (1), 774−784.
(83) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.
(84) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **2012**, *5* (1), No. 367.
(85) Hockney, R. W.; Goel, S. P.; Eastwood, J. W. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **1974**, *14* (2), 148−158.
(86) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126* (1), No. 014101.
(87) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.
(88) Darden, T.; York, D. M.; Pedersen, L. G. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(89) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463−1472.
(90) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13* (8), 952−962.
(91) Morra, G.; Potestio, R.; Micheletti, C.; Colombo, G. *Distance Fluctuation (DF) Analysis*. https://github.com/colombolab/Distance-Fluctuation-DF-Analysis (accessed 2023−12−05).
(92) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084−3095.
(93) Cock, P. J. A.; Antao, T.; Chang, J. T.; Chapman, B. A.; Cox, C. J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; de Hoon, M. J. L. Biopython: freely available Python tools for

computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25* (11), 1422−1423.
(94) Rehn, A.; Moroni, E.; Zierer, B. K.; Tippel, F.; Morra, G.; John, C.; Richter, K.; Colombo, G.; Buchner, J. Allosteric Regulation Points Control the Conformational Dynamics of the Molecular Chaperone Hsp90. *J. Mol. Biol.* **2016**, *428* (22), 4559−4571.
(95) Morra, G.; Neves, M. A. C.; Plescia, C. J.; Tsustsumi, S.; Neckers, L.; Verkhivker, G.; Altieri, D. C.; Colombo, G. Dynamics-Based Discovery of Allosteric Inhibitors: Selection of New Ligands for the C-terminal Domain of Hsp90. *J. Chem. Theory Comput.* **2010**, *6* (9), 2978−2989.
(96) Li, H.; Chang, Y.-Y.; Yang, L.-W.; Bahar, I. iGNM 2.0: the Gaussian network model database for biomolecular structural dynamics. *Nucleic Acids Res.* **2016**, *44* (D1), D415−D422. (accessed 2023−07−13)
(97) Bahar, I.; Lezon, T. R.; Bakan, A.; Shrivastava, I. H. Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins. *Chem. Rev.* **2010**, *110* (3), 1463−1497.
(98) Yao, X.-Q.; Hamelberg, D. Detecting Functional Dynamics in Proteins with Comparative Perturbed-Ensembles Analysis. *Acc. Chem. Res.* **2019**, *52* (12), 3455−3464.
(99) Pandini, A.; Fornili, A.; Fraternali, F.; Kleinjung, J. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. *FASEB J.* **2012**, *26* (2), 868−881.

https://doi.org/10.1021/jacs.3c11396
J. Am. Chem. Soc. 2024, 146, 901−919