Decrypting the languages of allostery in membrane-bound K-Ras4B using four complementary *in silico* approaches

Matteo Castelli,[1] *Filippo Marchetti,*[1,2] *Stefano A. Serapian,*[1,*] *Giorgio Colombo*[1,*]

[1]Department of Chemistry, University of Pavia, viale T. Taramelli 12, 27100 Pavia, ITALY.

[2]INSTM, via G. Giusti 9, 50121 Florence, ITALY.

[3]E4 Computer Engineering, via Martiri delle libertà 66, 42019 Scandiano (RE), ITALY.

stefanoartin.serapian@unipv.it; g.colombo@unipv.it

K-Ras4B, allostery, shortest path map, distance fluctuation, molecular dynamics, D-NEMD.

Abstract: Proteins are evolved molecular machines whose diverse biochemical functions are (also) dynamically regulated by allostery, through which even faraway protein residues can conformationally communicate. Allostery can express itself in different ways, akin to different "languages": allosteric control pathways predominating in an unperturbed protein are superseded by others as soon as a perturbation arises—*e.g*, mutation—that alters its function (pathologically or not). Accurately modeling these often-unintuitive phenomena could therefore help explain functional changes in a specific protein whenever they are unclear. Unbiased molecular dynamics (MD) simulations are a possibility; however, since allostery can operate at longer timescales than

those accessible by MD, simulations require integration with a reliable method able to, *e.g.*, detect regions of incipient allosteric change, or likely perturbation pathways. Several such (valid) methods exist, but are typically applied singularly, only yielding a partial ("monolinguistic") picture of allostery: we argue their joint application could significantly enrich this picture. To prove this, we perform unbiased MD simulations (~100 μs) of the widely studied, allosterically active oncotarget K-Ras4B, solvated and embedded in a phospholipid membrane, proceeding to decrypt its allostery using four showcase "languages": *Distance Fluctuation* analysis and the *Shortest Path Map* capture allosteric communication hotspots at equilibrium; *Anisotropic Thermal Diffusion* and *Dynamic Non-Equilibrium MD* assess them once the GTP that oncogenically activates K-Ras4B is, respectively, either superheated or hydrolyzed. "Languages" provide a uniquely articulate, mutually coherent, experimentally consistent picture of allostery in K-Ras4B. At equilibrium, pathways stretch from the membrane-embedded hypervariable region all the way to the active site, touching known flexible allosteric "switches" and proposed pockets. Upon GTP cleavage/perturbation, known to inactivate K-Ras4B, allosteric signals most reverberate on switches and interfaces that recruit effector proteins. Our work highlights the benefits of integrating different allostery detection methods, even in conjunction with unbiased MD.

Introduction

Proteins are efficient and versatile machines morphed by billions of years of evolution to oversee a variety of vital biochemical processes and support cellular life.[1,2] To meet these requirements, proteins populate a wide and diverse array of structures that are intrinsically dynamic[3-5] and are required to sustain well defined and finely tuned motions to carry out their functions[6-8].

Allostery[6,9,10] is a fundamental phenomenon 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 in constant expansion since seminal studies in the 1960s[18] and 1980s,[19] with the topic being extensively reviewed and updated in its description.[4,7-10,20-24]

Dynamic (residue-mediated) allosteric communication pathways arise even in the more conformationally constrained proteins,[6-8,19] favoring the transition between distinct energy-wells in biomolecules characterized by well-defined "folding funnel" minima.[5,10,23,25] Such minima (at the bottom of funnels) associated with the native structure can in fact be rugged and different sub-states can exist as rapidly interconverting conformational ensembles of slightly different functional/nonfunctional forms. Allostery is the key factor regulating kinetic transitions between these forms so that they coexist in precisely the right proportions required to ensure biological functions. In this framework, a targeted allosteric 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); or an endogenous ligand binding at an allosteric site; or upon complexation with another protein. In this way, conformational equilibria are subtly altered resulting in a "population shift"[4,7,8,10,20,21,23,25-27] and, with it, allosteric signals are relayed for a new biochemical event to occur, not necessarily in the vicinity of the perturbation site(s). These events can result in the modification of the properties of a particular interface (often distal to the perturbation site), with consequent promotion or disruption of another protein's recognition;[28,29] regulation of enzymatic reactivity[13] or the turnover rate of a molecular machine;[30] substrate binding or release;[31] and facilitation of conformational change.[14]

With such a delicate set of conformational equilibria required for normal biological functions, it is unsurprising that aberrant allosteric perturbations—whether introduced by PTM(s) or detrimental residue mutation(s)—are enough to disrupt these equilibria and lead to the insurgence of a number of pathologies.[22] Indeed, a promising therapeutic strategy[21,25,29] against such diseases, is to design small-molecule allosteric modulators[25,28,29] that bind to and interfere with identified allosteric pockets, providing a possible alternative to ineffective or toxic *ortho*steric ligands.[21,25,28] To provide another example in a different context, the concept of allosteric tweaking also plays a fundamental role in the field of biocatalysis,[8,11,12,26,31,32] whereby it is a major objective to reliably predict which residues in an enzyme (far from the active site) could be purposely mutated to obtain a new enzyme[11,12] with a specific, desired change in reactivity.

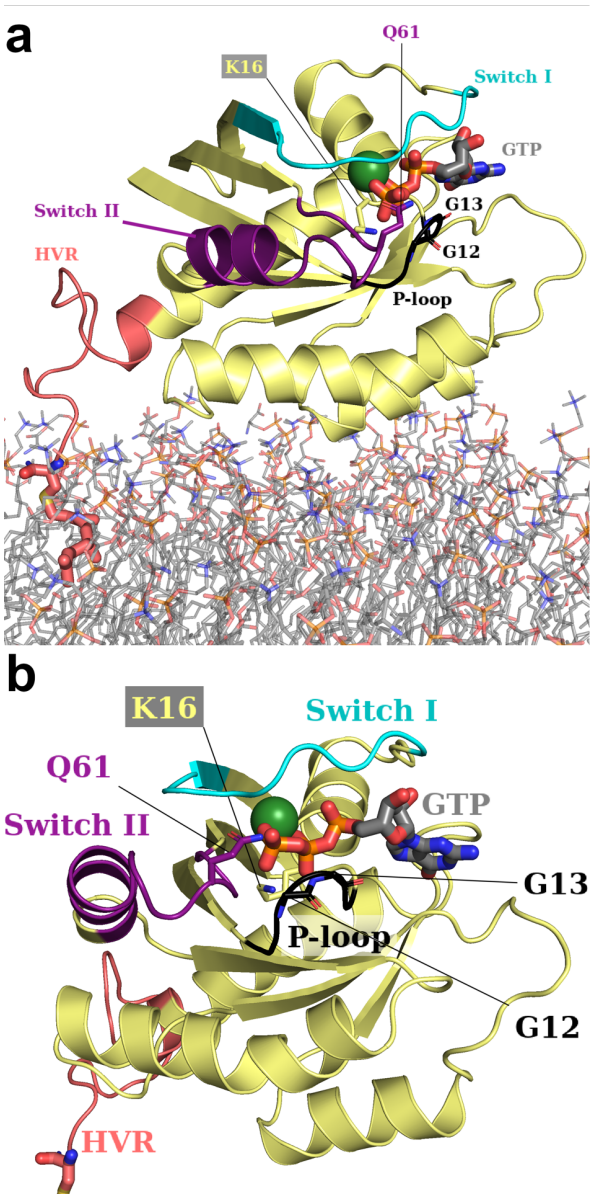

Figure 1. Structural overview of active K-Ras4B as simulated in this work. (a) Zoom on K-Ras4B as it appears in our starting structure (Figure S1a). Modelled after after PDB ID: 6vjj,[33] the G domain (residues 2-166; yellow) is represented with its salient features (see main text): switches I and II (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.[34] (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. In both panels, GTP is represented as thicker sticks; Mg^{2+} as a dark green sphere (both originate from PDB ID: 6vjj).[33] H, Na^{+}, Cl^{-} and solvent are omitted for clarity. Color code for explicit non-C atoms/ions: Mg^{2+}: dark green; Cl: pale green; Na: purple; P: orange; O: red; N: blue; S: yellow.

The signal-transducing kinase K-Ras[35,36] in its most abundant oncogenic isoform K-Ras4B[35] (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 1; salmon) terminating with a farnesylated Cys185 that is responsible for its incorporation into the cellular membrane (Figure 1a).[34,36,37] Under healthy conditions, membrane-bound K-Ras cycles between an *active* (GTP-/Mg^{2+}-bound) state (Figure 1) and an *inactive* one wherein GTP has hydrolyzed to GDP. When K-Ras is active, key regions of its G domain (switches I and II; Figure 1)[36,38] can morph allosterically to recruit and help activate various effectors[35,36,38] which then trigger appropriate signaling cascades. Subsequent K-Ras deactivation through GTP hydrolysis also requires switches I and II to adopt specific conformations[36,38,39] and, while intrinsic to a degree,[38] it is greatly facilitated[38,39] by the recruitment of a GTPase-activating protein (GAP) which immobilizes a catalytically crucial[39] Gln61 in K-Ras and administers an equally crucial[39] arginine (Figure S1b). Hydrolysis results in an inactive G domain with switches allosterically incapacitated to recruit effectors. In as many as 1 in 10 cancers[40] K-Ras4B is seen to undergo a plethora of different mutations[35,36,40] and/or PTMs[36] that, in ways that are not always allosterically clear[36,40] (bar steric

disruption of the K-Ras4B–GAP interface or catalytic interference, *e.g.*, through mutation of key residues Lys16 or Gln61; Figure S1 and Figure 1),[39] hinder GTP cleavage and thus trap the kinase in a harmful *hyperactive* state.

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 common mutants (*e.g.*, [35,40-42]), culminating in the recent (titanic) experimental effort by Weng *et al*.[40] to produce a comprehensive allosteric map of the effects of K-Ras mutations at all possible positions along its sequence. Significantly for this work, K-Ras4B dynamics and their variations in mutants have also been intensely studied from a computational point of view,[34,36,37,41,43-52] providing crucial insights on allostery and allosteric effector recruitment.[51,52] For a complete picture we certainly refer the reader to Pantsar's comprehensive review[36] but, for example, molecular dynamics (MD) simulations have confirmed: (i) that G domain switches are flexible[45,51] and flexibility is altered in different activation states[45,51] and oncogenic mutants;[48] (ii) that the G domain can fold back on the HVR and is affected by it,[43] even when K-Ras4B is embedded in the membrane;[34,44] (iii) that this interaction is weakened in oncogenic mutants;[47] (iv) that the kinase can rotate when anchored to the membrane;[34] and that (v) there exist several allosteric pockets in the G domain[41,51] (later explicitly assessed in the aforementioned experimental allosteric map).[40] Yet, despite all these findings and the clear medical need, covalent allosteric inhibitor Sotorasib (AMG510),[53] which rescues GTP cleavage in the oft-recurring K-Ras4B mutant G12C (*cf*. G12 in Figure 1), remains the only clinically approved option so far;[40] a noncovalent pan-mutant inhibitor BI-2865, impeding restoration of the GTP-active state by preventing nucleotide exchange, was reported as late as April 2023.[54] Such paucity of medical options is indicative of the infamously difficult druggability of K-Ras4B and the sheer allosteric diversity of

its oncogenic aberrations,[35,36,40] justifying the continued interest in the kinase and its allosteric mechanisms.

Computational techniques[21,23] are often essential to obtain an atomistic understanding of the determinants of allosteric regulation at levels of detail that remain inaccessible to experimental studies.[21,23,24] While the timescales accessible by atomistic simulations (now typically on the microsecond scale) remain on the short side when it comes to capturing complete conformational changes associated with allostery,[27] a number of solutions, including coarse-graining and enhanced sampling MD, have become available over the years (*cf.* [21,23,27,55]), also in combination with machine learning.[56] Markov State Model (MSM) analysis, Replica Exchange MD, and Principal Component Analysis have also been applied specifically to K-Ras4B.[45,46,48,51]

In this context, there exist atomistic-resolution methods rooted on unbiased MD that, while starting from different initial assumptions, aim at tackling the same question on unveiling the mechanistic details of allosteric modulation and the key residues involved; this is reminiscent of how different languages worldwide may be used to express identical concepts.

Here, exploiting the considerable trove of allosteric information on K-Ras4B and the interest it continues to attract, we assess the performance of four such techniques on (4.9 μs long) atomistic MD simulations of K-Ras4B, using them in a synergistic fashion. Our aim is to extract consensus information pointing to the key residues and dynamic traits defining 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" routes in large proteins and complexes thereof (*e.g.*, [14,16,57]); (ii) the shortest path map (SPM) method devised by Osuna and coworkers[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) technique initially

developed by Ciccotti and Jacucci,[58] which maps allosteric signal transduction from effector and/or substrate binding sites in receptors and enzymes;[17,59-63] and (iv) anisotropic thermal diffusion (ATD)[64] which, similarly to D-NEMD, entails heating an effector in its binding site within a (supercooled) protein or complex, and monitoring whereto the allosteric message propagates. Starting from the recapitulation of the allosteric intricacies of K-Ras4B in relation to the most recent findings, our goal is to offer a view of how often seemingly independent approaches, developed from different perspectives, speak in reality 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.

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 serve as the starting point for further nonequilibrium MD simulations *per* the final two allostery detection methods (languages), *i.e.*, dynamical nonequilibrium MD (D-NEMD)[17] and anisotropic thermal diffusion (ATD).[64]

System Setup. All equilibrium MD simulations were begun from a single representative structure (Figure S1a) of mature, active K-Ras4B embedded in a negatively charged and previously equilibrated[34] phospholipid bilayer 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.[36]

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 1).[33] From this structure we proceeded with the *PyMol* package[65] as follows: we removed the co-crystallized 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[39] conformation, the Gln61 sidechain was rotated inwards to match its observed orientation in the presence of a GAP (Figure S1b; *cf.* PDB ID: 1wq1);[66] 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 coworkers[34] 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)[33] 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;[33] we note that the G domain in 6vjj deviates little from the simulated one (RMSD 0.785 Å), and we comment on the differences in the *Results and Discussion* section. At this stage, for simplicity, we also capped the *C*-terminal Fcy185 with an *N*-methyl group (rather than the *O*-methyl group that is present in the mature form).[36]

After superimposing and merging, the *reduce* utility in *AmberTools* (v. 19)[67] was employed to add hydrogens to K-Ras4B G domain residues; predict histidine tautomerization (on Nε2 in all cases); model optimal orientations of Asn/Gln sidechains (ignoring the aforementioned catalytic Gln61); and confirm the absence of disulfide bridges. The *PropKa* package[68] predicted all residues to be in their standard protonation states at physiological pH. Finally, the *tleap* utility[67] 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[34] (*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*[67] was converted to a *GROMACS*-compatible[69] format using the *pdb2gmx* utility: starting coordinates are available as Supporting Information.

Forcefield Parameters. For reasons of mutual compatibility with other parameters, all standard K-Ras4B amino acids and terminal caps were modeled using the *ff99SB* forcefield[70] in its *ILDN*

improvement,[71] whereas parameters for Fcy185 were based on the work by Khoury *et al*.[72] With regard to ions, Na^+ and Cl^- were treated with parameters by Joung and Cheatham,[73] while parameters by Allnér and coworkers[74] were used to model Mg^{2+}. For GTP and GDP (the latter present in D-NEMD simulations only; *vide infra*), we adopted the forcefield reported by Meagher *et al*.[75] Similarly, to simulate the inorganic phosphate anion $[H_2PO_4]^-$ present in D-NEMD simulations only we introduced parameters by Kashefolgheta and Vila-Verde.[76] Parameters for both lipids present in the membrane (POPC and POPS) were provided by appropriate extensions of *Slipids* forcefield[77,78] by Jämbeck and Lyubartsev. Finally, the chosen water model was TIP3P.[79] Where necessary, parameters were converted to *GROMACS*-compatible[69] formats using the *acpype* code,[80] 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*[70,71] from forcefields of the Charmm family originally used by Prakash and coworkers,[34] no major structural upsets were observed in the equilibration stages of the MD simulations (*cf*. next subsection): this confirms that the change of forcefield 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),[69] 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 timestep—lengthened from the preproduction stages, see Supporting Information—applied to the leap-frog integrator.[81] The 300 K temperature was enforced by the velocity-rescaling thermostat,[82] to which protein, membrane, and solute/solvent were coupled separately with a 100 fs time constant. To enforce the 1 bar pressure during production, we applied Berendsen's barostat[83] 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 for the calculation of Lennard-Jones and Coulomb interactions, and for determining the closest neighbors around each atom (lists were updated every 10 integrator steps). Beyond 12 Å, Coulomb interactions were calculated with the *Particle Mesh Ewald* method[84] rather than in direct space, whereas Lennard-Jones interactions were set to zero: effects of this were compensated, as *per GROMACS* implementation, by adding average corrections to the energy and pressure. All non-water bonds containing hydrogens were constrained with the *LINCS* algorithm;[85] water bonds were constrained using *SETTLE*.[86] All unspecified details were set to *GROMACS* defaults;[69] all input files are provided electronically as Supporting Information.

Given the change in parameters and release of constraints when moving from equilibration to production, we verified that it took at most 5 ns before each replica truly equilibrated. We thus only considered the last 245 ns of each replica as the true "production" stage: as such, for all ensuing analyses on allostery and to set up (nonequilibrium) simulations, we concatenated the last 245 ns of each of the 20 replicas into a single metatrajectory (total: 4.9 μs).

Distance Fluctuation Analysis (DF). Broadly speaking, distance fluctuation (DF) analysis[14,16,57] 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 ith 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 ith 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 in a more uncoordinated way and are less likely to be allosterically related; *vice versa*, residue pairs exhibiting low DF scores are moving concertedly and are deemed to be allosterically connected.

DF analsysis on equilibrium MD simulations of K-Ras4B in this work was conducted using our *ad hoc* code,[87] directly on our 4.9 μs metatrajectory; the procedure does not require any fitting or realignment.

Shortest Path Map (SPM). The theory behind SPMs (shortest path maps) 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 ith and j^{th} residues) are 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 is found through

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

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 except its sequential neighbors: to give an idea, K-Ras4B, with its 184 residues, features 16652 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 4.9 μs metatrajectory (in .xtc format) using the "matrix" command in the *cpptraj* tool,[88] after a few trivial realignment steps and a clustering procedure based on Cα atoms: these steps are detailed in Supporting Information. We additionally repeat the analysis, for comparison, on the full 5 μs

metatrajectory (*i.e.*, considering the production stage as a whole). As Supporting Information, we also provide 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*, for both the 4.9 μs and 5 μs analyses.

Dynamical Nonequilibrium Molecular Dynamics (D-NEMD). Short D-NEMD simulations (6 ps rapid equilibration + 44 ps production when surviving to completion; *vide infra* and Figure 2) were conducted starting from 71706 individual frames ("windows") directly isolated from our 4.9 μs metatrajectory with atomic velocities (stored every 50 ps; Figure 2, 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 chemical characteristics previously[39] established *in silico*, including the presence of a nucleophilic water in the vicinity of the γ-phosphate in GTP (criteria and exact "reactive" frame counts *per* replica are provided as Supporting Information).

The concept behind D-NEMD, which is extensively reviewed by Oliveira *et al.*,[17] is illustrated in Figure 2: an identical perturbation is instantaneously introduced in each isolated window, after which MD is resumed for a short period, and regions of greater or smaller deviation from equilibrium MD (*i.e.*, how perturbative effects propagate) are monitored frame-by-frame and 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 forcefield-dictated[75,76] equilibrium geometries. This charge-preserving interconversion is achieved by reordering and shifting the positions of as few as four atoms (details and illustration in Supporting Information), without altering any of the atomic velocities issuing from equilibrium MD.

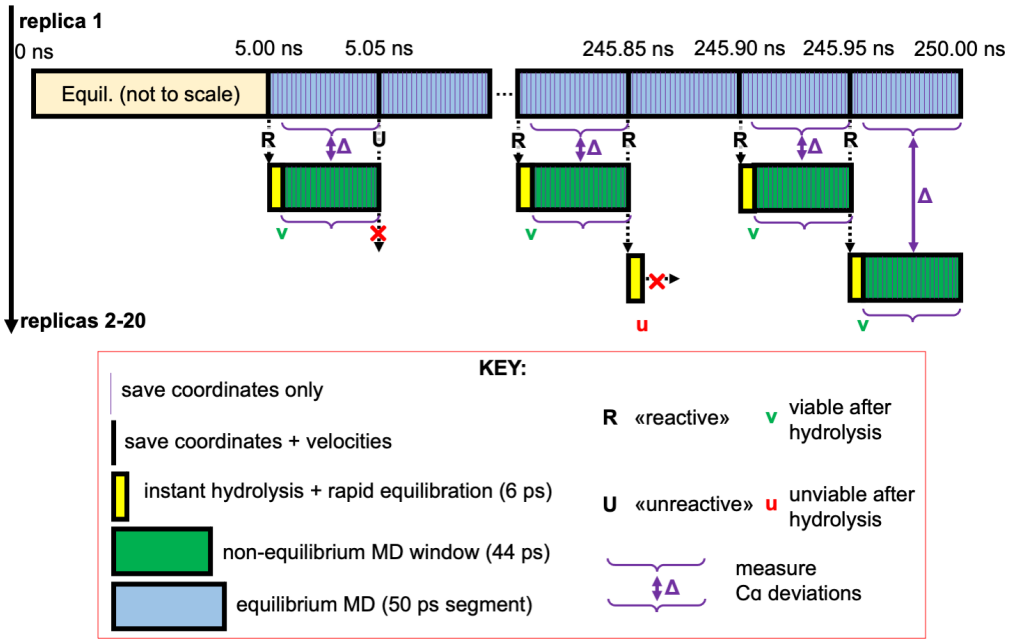

Figure 2. Overview of D-NEMD simulations setup, exemplified for equilibrium MD replica 1 and repeated for the remaining 19 replicas. Barring an initial 5 ns equilibration period (pale orange box), poses saved with velocities every 50 ps in the 5 to 245.95 ns interval (thick black lines) are assessed based on loose "reactive" (R) or "unreactive" (U) character (see main text). From each "R" pose, a new D-NEMD window is initiated after instant hydrolysis of GTP (see main text and Supporting Information). After rapid equilibration (6 ps; yellow boxes), windows developing unviable product geometry (u) are discarded, while D-NEMD production (44 ps; green boxes) is initiated for viable geometries (v). Deviation of Cα atoms between equilibrium and non-equilibrium MD is then monitored across each window pair, and averaged across all windows and replicas to obtain final per-residue values at regular intervals over 50 ps.

Starting from this new topology (provided electronically as Supporting Information), each window is first made to equilibrate very rapidly: over 6 ps, with an ultra-short timestep of 0.6 fs, and *LINCS*[85] and *SETTLE*[86] 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 equilibration (the verification process is detailed in Supporting Information),[89] we then restore the same conditions present in equilibrium MD and continue simulating each window for a further 44 ps, reaching 50 ps (Figure 2; 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.

As hinted, both the D-NEMD production and rapid equilibration stages—the latter indicated as "slow growth" in previous literature,[59,60] where it was applied by gradually switching on the new charges—come at a loss, with 80.6% of the initial "reactive" windows surviving to completion. This low loss rate, which is due unphysical breakup of the reconstructed $[H_2PO_4]^-$ in certain 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 × 57823 windows). Exact totals are provided in Table S1.

Anisotropic Thermal Diffusion (ATD). ATD MD simulations were conducted starting from a set of 4900 frames isolated from the 4.9 μs metatrajectory (*i.e.*, at 1 ns intervals). 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 retained at 12 Å; 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 2ps. 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; input provided as Supporting Information).[71] 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 some contextualization and comparison to existing computational and experimental understanding of K-Ras4B. There finally follows a brief critical discussion on the implications of our choices and their findings.

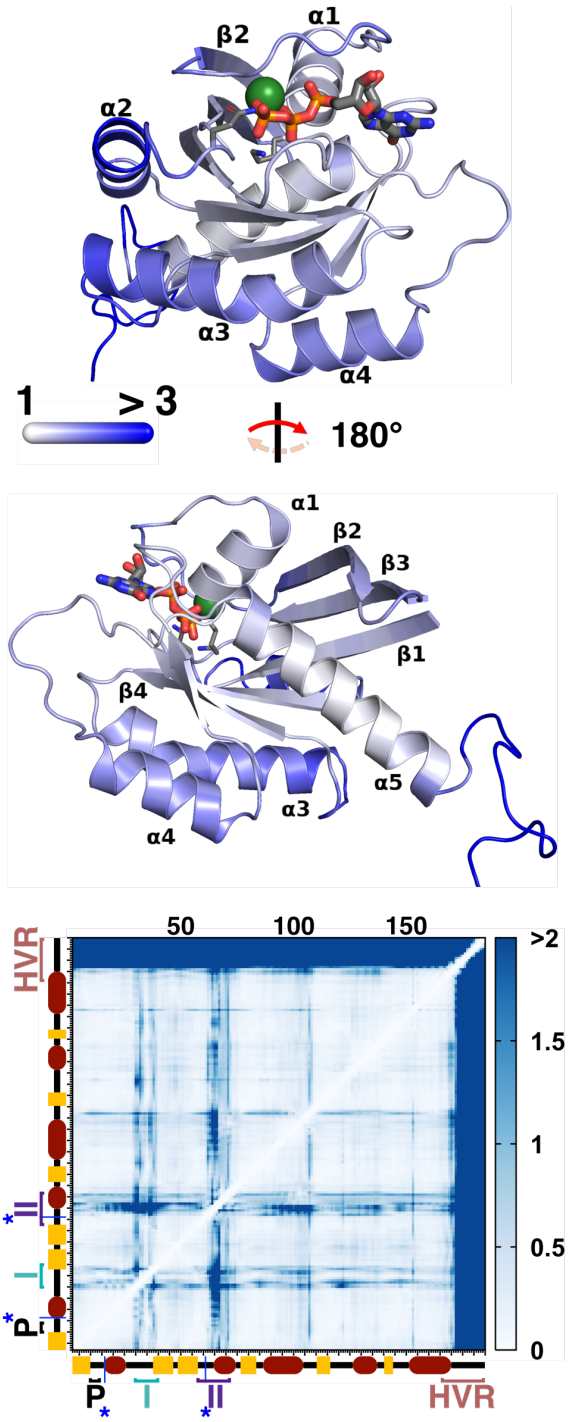

Figure 3. (top) Projection onto the starting K-Ras4B structure (Figure 1a) of the average (1D) DF score "felt" by each residue (see text for details), with backbone colored in increasingly darker shades of blue for increasingly higher DF scores. (middle) The same structure, rotated 180° about

the axis parallel to the plane of the page. In both views: GTP is rendered as ticker 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, secondary structure is labeled explicitly. (bottom) 2D matrix of pairwise DF scores. 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).

Distance Fluctuation Analysis on Equilibrium MD. As a reminder, *higher* DF scores denote residue pairs moving *less* concertedly 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 3; bottom; with secondary structure elements marked along its axes): this 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 unstructured regions exhibiting significantly low coordination to the rest of the G domain include the unstructured part of switch II just downstream of catalytic[39] Gln61; the *N*- and *C*-termini of the unstructured part of switch I (residues 30-33 and 36-37); and residue 107. Remaining unstructured areas are instead tendentially more coordinated (*cf.* less intense tones of blue in the matrix): such regions notably include central switch I residues 34-35 (the latter a Mg^{2+}-bound Thr that unbinds upon switch I deactivation);[42] catalytically relevant Lys16 (and to a lesser extent, Gln61); as well as the mutation-prone P-loop, whose only prominent coordination losses are with respect to the HVR, switch II, and (albeit less so) switch I.

Turning to regions of secondary structure, it can be seen that β-sheets are associated to (mostly) "white" areas in the matrix, suggesting they may form a coordinated and compact hub, with a high degree of inter-residue crosstalk that is crucial for the repartition of allosteric signals to other K-Ras4B regions. For the best part, helices α5 and α1 (within which Lys16 is contained) also exhibit high coordination to most G domain areas; helices α3 and α4 follow suit, with slightly "bluer" columns in the matrix and therefore a somewhat poorer generalized coordination. Conversely, likely owing to its significant overlap with the second half of switch II (over residues 66-72), helix α2 emerges as one clear outlier, giving rise to prominent blue areas in the DF matrix, exactly like part of the unstructured half of switch II: this signals very low coordination with the rest of K-Ras4B. Another allosterically uncoupled exception are the three *N*-terminal residues in β2, comprised in the *C*-terminal part of switch I. A final outlier is residue 153, also appearing in the matrix with a distinctive blue band despite its position at the *N*-terminus of (very rigid) helix α5.

To aid in the interpretation of the DF matrix in Figure 3, 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 accordingly.

In addition to this video, in the top part of Figure 3, we have projected a 1D-averaged version of the DF matrix onto our starting K-Ras4B structure (*cf*. Figure 1), which is also represented just beneath it (Figure 3; 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, 2D information contained in the matrix is lost, but one gains insight into the "average" degree of allosteric (un)coordination experienced by each residue.

Besides clearly recapitulating the elevated rigidity of most of helices α1 (including Lys16) and α5 (Figure 3; middle; white), the moderate rigidity of the P-loop (Figure 3; top; bluish-white), and the prominent decoupling of most of switch I, most of switch II, and the HVR (Figure 3; top; blue), the 1D projection brings to light certain traits that remain hidden in the 2D data. In particular, helices α3 and α4 become bluer (less coordinated on average) than they appear in the 2D matrix, where their loss of crosstalk is all concentrated with respect to switch II. Similarly, while a greater coordination within β-sheets is still recognized, sheets β4, β5, and β6 (best visible in the middle structure in Figure 3) are on average more allosterically interlinked compared to β1 and (especially) β3 and the part of β2 that isn't comprised in switch I: in the case of the latter two, in light of their acquisition of a bluer color, it can be concluded that on average they are less coordinated than initially inferable by inspecting the 2D matrix.

To summarize, 1D and 2D DF data alike concur in finding sheets β4–β6 to be particularly coordinated in their movements, and thus exhibiting high allosteric intercommunication; the same rigidity is observed for helix α5. Most of the G domain only retains a moderate degree of internal coordination, with the P-loop; possibly sheets β1, β3 and most of β2; catalytic Lys16 as part of helix α1; and (to a lesser extent) Gln61 and very minor portions of switches I and II falling in this category. In stark contrast, on the other hand, 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 clearly uncoordinated movements.

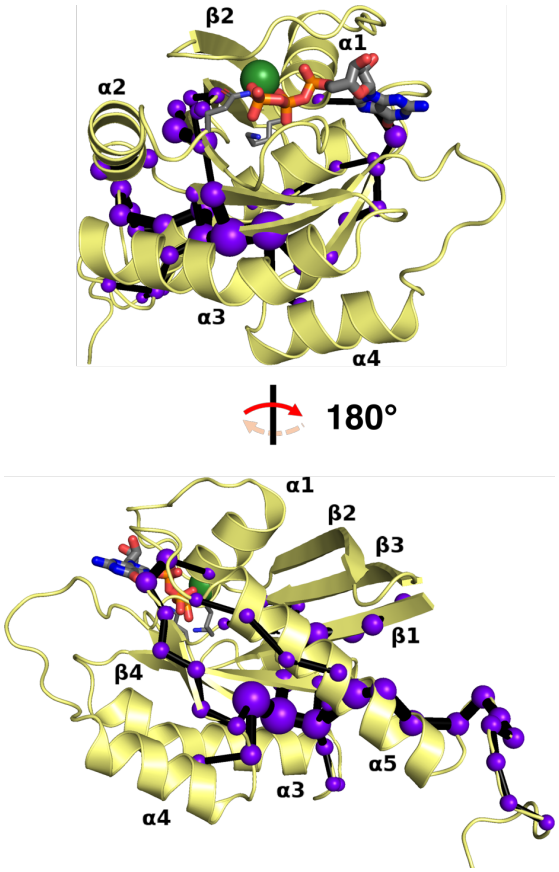

Figure 4. Allosteric shortest path map calculated from our 20 concatenated 245 ns equilibrium MD simulations of K-Ras4B. The map, wherein purple spheres represent communicating residues (Cα atoms) and thickness of black sticks is proportional to their "allosteric closeness" (see text), is drawn on the starting structure; views and atom color codes are identical to those for Figure 3.

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 4.9 μs metatrajectory, is illustrated in Figure 4 as purple spheres (residues) connected by black lines (paths), which we project onto the same (starting) structure present in Figure 1, also rotated by 180° as in Figure 3 (middle). The electronic *PyMOL* session file used to derive Figure 4 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 4, and proportional to the intensity of allosteric communication between the two vicinal residues it connects; the reader should also bear in mind that only segments with a relative shortness of 0.3 out of 1 or above are arbitrarily represented in the SPM, meaning that residue pairs whose coupling is just below this limit will not show up, and portions of the SPM will occasionally appear as isolated fragments.

Recalling the above, interpretation of SPM results is more straightforward than DF analysis, though the two paint a very similar picture as commented later in this subsection. Indeed, the SPM features a prominent allosteric pathway that originates midway along the HVR (Figure 4; 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 α5 helix, the path first reaches Ile163; Ile163 communicates very strongly with residues Val112, Met111 at the *N*-terminus of sheet β5, and Pro110 on the α3–β5 loop, with an elevated relative shortness of 0.79 out of 1 (Figure 4; bottom; *cf. Computational Methods*). Indeed, by virtue of this shortness, Ile163 and the three α3–β5/β5 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 kinase (shortness: 1 out of 1).

Five shortest path branches depart from this four-residue hub. Branch I (Figure 4; bottom) returns to Glu162 from Pro110 and conveys allosteric signals along the almost entirety of helix α5, incidentally avoiding the flexible Asp153 identified by DF analysis. Branch II (Figure 4; bottom) reverberates to sheet β6, through which it travels in its entirety before reaching Ser145 and Ala146 on the β6–α5 loop and therefore the binding site (guanine moiety), and further jumping across to Leu19 on helix α1. Branch III (Figure 4; bottom) terminates abruptly at Ala134 in the *C*-terminus

of α4, after crossing to the α4–β6 loop. Branch IV (Figure 4; bottom) departs Pro110 and moves back towards the *N*-terminus along the α3–β5 loop, but dissipates at Ser106 before it can reach helix α3 (Lys104) or the hyperflexible Glu107.

From an allosteric point of view, Branch V is the most intensely travelled route and the one spanning the most interesting regions. It branches from the hub on β5 off to β4, with a dual coupling to Gly77 from Pro110 (shortness: 0.42 out of 1) and Met111 to Phe78 (0.64 out of 1); it then forms one subbranch V.1 crossing almost the entirety of sheet β1 and one V.2 just reaching the *C*-terminal Arg73 on α2, just outside switch II. A third subbranch V.3 touches Val7 on β1 and the *C*-terminal residues of β3, before moving (Figure 4; top) all the way up to the (energetically important)[39] active site residue Ala59 on the *N*-terminus of switch II, stopping just short of Gln61. In fact, when rederiving the SPM on the 5 μs metatrajectory as a control (*i.e.*, including rather than excluding the first 5 ns in each replica; Figure S2; *PyMOL* session also provided electronically), we were intrigued to find that the only major difference was actually the extension of subbranch V.3 to oft-mutating P-loop glycines 12 and 13, and Mg^{2+}-coordinated Thr35 on Switch I.

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. A few areas within the binding site or its vicinity are also eventually reached: Thr58 and Ala59 on switch II; Thr35 on switch I (only if the 5 μs metatrajectory is considered); and Leu19 on helix α1. The path also marginally touches sheet β3 and helices α2 and α4, whereas helix α3, the remainder of switches I and II (minus Thr58 and Ala59), and the remainder of sheet β2, remain out of reach. Like Thr35 on switch I, P-loop represents an interesting case: its appearance in the SPM when calculated over 5 μs but not in that calculated over 4.9 μs suggests a moderate but

nonnegligible allosteric importance, precisely as found by DF analysis. Indeed, consistency of SPM findings with DF data is readily ascertainable, particularly in terms of the 1D average DF scores projected on the structure in the top and middle of Figure 3. 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 3). *Vice versa,* areas with low average coordination coincide with those untouched by the SPM.

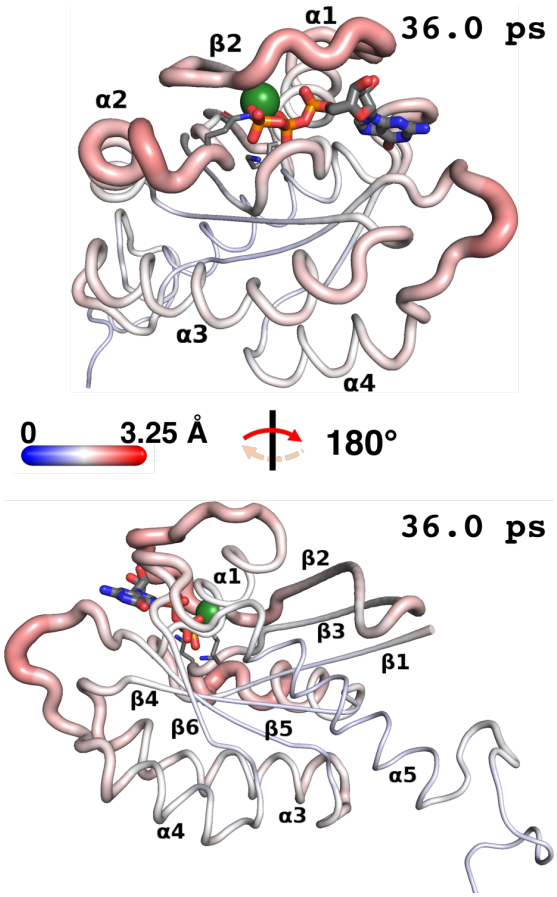

Figure 5. D-NEMD simulation results viewed on the starting K-Ras4B structure (views and atom color codes are identical to those for Figure 3). Backbone thickness of each residue and intensity of red *vs*. blue is proportional to the average absolute deviation of its Cα atom in D-NEMD simulations—*i.e*., 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; in this Figure they are exemplified as they appear 36.0 ps after hydrolysis, but they were measured up to 50.0 ps after. The choice of time is the one that visually maximizes contrast between most- and least-deviating residues at the chosen color scale (0 to 3.25 Å deviation); note however, that no residue reaches either of these values at the chosen time.

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 5), average perturbations resulting in D-NEMD simulations compared to equilibrium MD (greater in the redder areas) reveal 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 5), but not in a uniform way: it is clear (Figure 5; top) that switches I (2.04 Å average deviation at 36 ps) and II (1.93 Å) in their entirety bear a greater brunt. In fact, at 2.27 Å, Glu63 on switch II is the residue with the greatest average deviation compared to equilibrium MD, and catalytic Gln61 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 Å and 1.74 Å, respectively, make them appear much whiter in Figure 5 (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 by hydrolysis, deviating by only 1.57 Å.

Compared to equilibrium MD, core sheets β1 and β4–β6 and helix α5 (Figure 5; bottom) exhibit an even lower combined average deviation of 1.48 Å at 36 ps: 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 are themselves unaffected by hydrolysis. Focusing on remaining peripheral regions in K-Ras4B, it is equally clear from Figure 5 that not all of them are affected by D-NEMD hydrolysis in the same way: as in previous D-NEMD studies,[59-62] 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 D-NEMD. The only distal region from the hydrolyzing phosphate to undergo significant deviation, though still part of the binding site, is evidently the β5–α4 loop (on the right of Figure 5; top; and on the left of Figure 5; bottom), on a 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 5), but not to the same magnitude; these include, *e.g.*, the α1 helix (≤1.98 Å) in its farthest part from GTP, the β2–β3 loop (≤1.95 Å), the *N*- and *C*-termini of helix α3 (≤1.89 Å), the *N*-terminus of helix α4 (≤1.87 Å), and the β2 sheet outside switch I (≤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 (≤1.64 Å), entirely comparable to the α5/β1/β4–β6 core: this again suggests it is likely unaffected by allosteric signals emanating from the active site.

We should further remind the reader that D-NEMD statistics were collected on the 57823 loosely "reactive" equilibrium MD frames surviving D-NEMD hydrolysis, out of an original 71706: "reactive" character in these frames was loosely decided based on the detection of a nucleophilic water molecule in the vicinity of the Pγ atom of GTP (see precise criteria in Supporting Information); however, we chose not to consider the conformation of the Gln61 sidechain, which in reality is crucial to properly lock the nucleophile into position for attack.[39] Thus, for completeness, we proceeded to recollect statistics on only a subset of surviving D-NEMD frames in which the initial perturbed frame also had Gln61 in a suitable position: although this considerably reduced the dataset to 23209 surviving D-NEMD windows (out of 29400 "reactive" poses), we confirm that the outcome resulting from this reduced statistical set is virtually indistinguishable from the allosteric picture painted by deviations in Figure 5 (data are not shown, but are available upon request).

To recapitulate, D-NEMD 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 β6–α5 loop, while catalytic Lys16 is barely affected. Away from the active site, the rigid α5/β1/β4–β6 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 β5–α4 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 on reheating GTP alone to 300 K after supercooling the entire system at 10 K, as repeated starting

from 4900 individual frames isolated from the 4.9 μs equilibrium 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 feeling the effects of GTP hydrolysis. Also, unlike in D-NEMD, progression of allostery in ATD is measured with respect to the first production frame in each window (whence 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" on all 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 6).

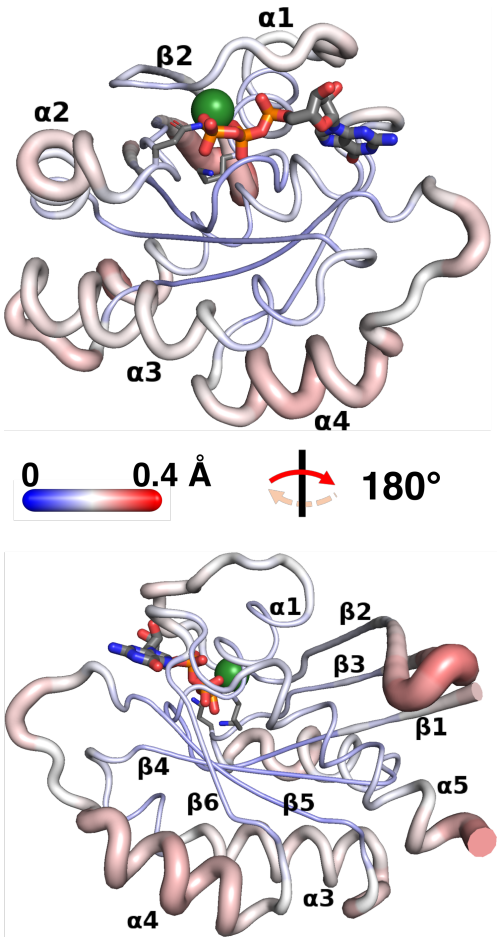

Figure 6. ATD results viewed on the starting K-Ras4B structure (views and atom color codes are identical to those for Figure 3; note, however, that in this case analysis was truncated beyond residue 168, and therefore most of the HVR is omitted). Backbone thickness of each residue and intensity of red *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.

ATD simulation data projected on the starting K-Ras4B structure (Figure 6) 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 5), 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 5), and in contrast to their uncoordinated (flexible) status detected in the SPM (Figure 4) and by DF analysis (Figure 3). 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 α2 (Figure 6; top): even visually, it is clear that there are areas in K-Ras4B that *per* Figure 6 show greater or comparable average deviation, all of which are far from the GRP binding site. The first of these is β2–β3 loop (Figure 6; bottom), whose Gly48 exhibits the greatest deviation of all residues considered (0.29 Å), followed by helix α4, with an average deviation of 0.24 Å. The α3–β5 loop, whose residue 107 stood out as allosterically uncoupled in the DF analysis, shows a comparable degree of average deviation (Figure 6; bottom), and the β5–α4 loop deviates just a little less, at 0.18 Å on average (Figure 6; top and bottom).

The only other region of the active site that shows appreciable deviation—which, incidentally, is found by D-NEMD too—is the β6–α5 loop, also at 0.18 Å (Figure 6; 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 6; top), and in catalytic residues Gln61 (Figure 6; top) and Lys16 (Figure 6; bottom), suggesting that while coupled to GTP once it is hydrolyzing (D-NEMD), these areas are not relevantly coupled to GTP.

Focusing, finally, on the least-deviating parts of K-Ras4B away from the active site (Figure 6), we once again confirm that the α5/β1/β4–β6 core is minimally perturbed, excluding the final few *C*-terminal residues of α5, 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 α5/β1/β4–β6 core is also found to be unaffected by GTP heating, but in this case much expectedly.

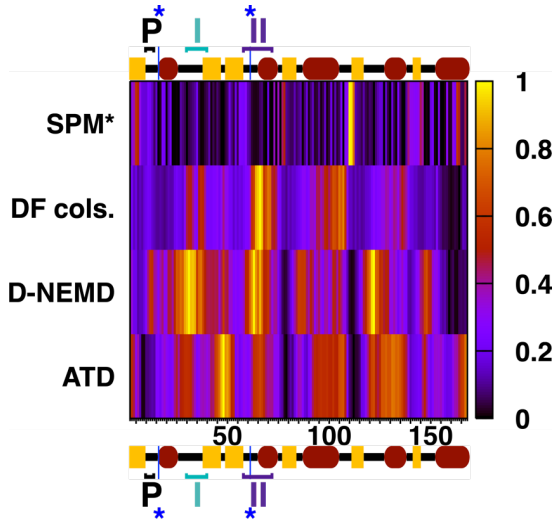

Figure 7. A more direct comparison of the four allosteric languages compared in this work, with per-residue representative scores for each method (see main text for chosen ones) normalized to between 0 (black) and 1 (yellow), passing through shades of purple, red, and orange. Since the ATD analysis excludes 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 ("betwenness") 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.

Listening to the Various Languages. While clear trends already emerge from isolated analyses in Figure 3–Figure 5, in order to obtain meaningful information about K-Ras4B it is of course necessary to compare the four allostery detection methods more systematically: this is possible from the heatmap plotted in Figure 7. In the heatmap, which only spans residues 2-168 since ATD was not meaningfully analyzable in the HVR, we compare a distinctive per-residue score S_{norm} chosen for each allosteric language, rescaled/normalized so that the score $S_{norm,i}$ for the ith residue will always fall between 0 (black) and 1 (yellow) as prescribed by the following formula:

$$S_{norm,i} = \frac{S_{raw,i} - S_{raw,min}}{S_{raw,max} - S_{raw,min}}$$

wherein $S_{raw,min}$ and $S_{raw,max}$ denote, respectively, the minimum and maximum raw scores detected for a particular method over the 168 residues considered. Essentially, therefore, the normalized score $S_{norm,i}$ measures how close the raw score for the ith residue $S_{raw,i}$ is to either $S_{raw,min}$ or $S_{raw,max}$ on a scale from 0 to 1, *i.e.*, it has a value of 0 for the lowest-scoring residue ($S_{raw,i} = S_{raw,min}$) and 1 for the highest-scoring residue ($S_{raw,i} = S_{raw,max}$), and falls in between these extremes for remaining residues.

Choice of raw scores S_{raw} in the case of ATD and D-NEMD (Figure 7; two bottom rows) was straightforward, since both already directly yield a per-residue value: ATD yields the average RMS deviation *per* residue upon commencement of heating, while D-NEMD calculates average absolute deviation of each Cα with respect to equilibrium MD. While DF scores produced by DF analysis are pairwise rather than per-residue, a natural choice for S_{raw} was to inherit and normalize (Figure

7; second row) the same averaged 1D DF scores that are projected onto the starting structure in the top and middle panels of Figure 3. Finally, the score chosen for SPM is directly obtainable from the *DynaComm.py* code and is termed *betweenness*:[12,31] it is proportional to the number of times that an individual residue gets 'visited' as an intermediate point along one of the many individual *shortest paths* that connect each residue to every other residue.

Normalized per-residue betweenness scores for the SPM are plotted on the top row of Figure 7, and we should stress that they are conceptually opposite to the other three chosen S_{norm} scores. In the former case, residues with the highest betweenness correspond by definition to those falling along the final SPM (Figure 4), and therefore S_{norm} scores tend to 1 in the regions with the highest allosteric communication/rigidity; conversely, in the case of the other three scores, residues for which S_{norm} scores tend to 1 are either those with *low* allosteric communication/rigidity or those with the greatest deviation upon heating or hydrolysis. For this reason, normalized SPM scores (Figure 7; top row) should be envisaged as following opposite chromatic trends with respect to the other three normalized scores.

Interpreting the Various Languages. With this *caveat* in mind, an overview of Figure 7 unequivocally indicates that there is a 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 (equally important) 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.

If allostery spoke a universal language, which is a tempting but unrealistic expectation, only one method, not four (or more, *vide infra*), would be enough to decipher all information about a system from its unbiased MD. Instead, it is precisely due to the different declensions of allostery that we have here chosen to employ more than one language, and we recall that this is the main argument driving our study. 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. Otherwise put, much like certain actual languages can sometimes express a concept more effectively or concisely than others through different syntactic rules or vocabulary, our chosen allosteric languages are sufficiently different in their theoretical framework so that they can adequately interrogate K-Ras4B on its propensity to visit allosteric states that become relevant at very different stages of its biochemical lifecycle.

More specifically, the SPM and the DF analysis are natural partners (akin to linguistic cognates), sharing several affinities. Unlike ATD and D-NEMD, they are true unbiased methods, really 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: either because their constituent residues are connected by shortest paths (Figure 4; *cf. Computational Methods*), or because their DF scores with respect to other parts of the kinase are particularly low. By virtue of their two-dimensional nature (Figure 3; 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 clearly does not undergo major conformational rearrangements at our

simulated timescales, hotspots uncovered by SPM and DF analysis represent areas that would be most likely to carry signals driving any such rearrangement at sufficiently long timescales, and that would bring about marked effects if interfered with in any way.

ATD and D-NEMD too 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: rather than explicitly detecting allosteric propagation routes, they identify the "endpoints" that ultimately experience the greatest repercussions from the allosteric perturbation in question. Regions whose dynamics don't change as much could certainly correspond to allosteric hotspots propagating the perturbation, but not necessarily: they may simply be uninvolved in conveying that particular perturbation, and this can only be deduced upon listening to one or more of the other languages. ATD is the more 'unphysical' language of the two: by artificially heating (but not hydrolyzing) GTP in a supercooled K-Ras4B, it reveals the allosteric "endpoints" of signals specifically emanating from the active site. In this respect, it could be for example integrated with a targeted analysis of DF scores between specific active site residues and the remainder of the kinase. D-NEMD similarly informs explicitly on allosteric "endpoints" rather than propagation routes; unlike ATD, however, it is based on instantaneous differences, and it furthermore monitors the result of a real biochemical event (phosphate cleavage), with tangible consequences. "Endpoints" identified by D-NEMD should thus be most directly relatable to biochemical consequences, *e.g.*, which are the distal areas that most prominently intercept signals originating from phosphate cleavage, and how changes in these areas could result in the creation or destruction of a biochemically functional interface.

The significance of the nuances characterizing the four allosteric languages will become better apparent in the subsections that follow, wherein we dissect normalized scores shown in Figure 7,

and focus first on situations when all languages are found to be in agreement, then on those when they are not. As crucially recounter later, all allosteric details emerging from our analysis are entirely in line with experiment and most recent allosteric understanding of K-Ras4B.[36,40,42,63]

When Linguistic Consensus Can Tell a Similar Story from Different Viewpoints. The most obvious consensus across all normalized scores in Figure 7 is that the highest allosteric coupling at equilibrium (SPM betweenness) and either some very low fluctuations (*i.e.*, DF) or susceptibility to perturbation (*i.e.*, D-NEMD, ATD) are invariably observed (yellow for SPM, black/darker purple for the other three S_{norm}) at the cusp of the α3-β5 loop and β5 sheet or, otherwise put, at the hub centered on residues 110-112 whose emergence from the data we have frequently referred to in preceding subsections: such a clear-cut consensus (*i.e.*, allosteric hotspots and not "endpoints") strongly proposes these residues as the main allosteric center in K-Ras4B. A very similar situation of consensus across languages is observed at the *N*-terminus of the β4 sheet, reconfirming that this sheet too is allosterically crucial, including for the propagation of allosteric perturbations. On the opposite front lie regions that are allosterically less coupled (more flexible) at equilibrium *per* SPM and DF, but at the same time clearly identified by ATD and D-NEMD as allosteric endpoints of perturbative events: these are the regions exhibiting low normalized SPM betweenness (darkest shades in Figure 7, top row) and some of the highest normalized DF/D-NEMD/ATD scores (lightest shades in Figure 7, bottom rows). Such zones include: the majority of switch I—though all four maps feature Thr35 as an outlier with a distinctly colored band; the central part of switch II; the short β2-β3 loop; and, less prominently, nonuniform patches of helices α3 and α4. Finally, there are intermediate areas, which due to uneven S_{norm} distribution tend to have all four Figure 7 scores in the 'purple' range: the main element in this category is the β6 sheet, which should be

envisaged both as retaining some degree of allosteric relevance at equilibrium, as well as crucially "receiving" some of the allosteric impulses originating from GTP hydrolysis or heating.

When Linguistic Nuances Can Narrate Allostery over Time. We here shift our focus on the equally fundamental areas of K-Ras4B wherefore S_{norm} values of one or more languages in Figure 7 utter different answers. Since there are very different implications depending on which methods give which answers, we won't discuss all these areas systematically; we will instead consider those areas that are most frequently considered in the literature, as well as secondary structure elements that weren't mentioned in the previous subsection (*i.e.*, those for which there is no consensus and at the same time aren't part of either switch).

Naturally, the first K-Ras4B area to discuss here is the mutation-prone P-loop: the discrepant language is ATD, whose S_{norm} signals a lower-than-expected deviation for this loop upon GTP heating (black band in Figure 7, bottom row; *cf.* also its blue color in Figure 6). Were it not for this single "linguistic outlier", the P-loop would resemble the β6 sheet discussed at the close of the previous subsection, with moderate contributions to allosteric signal transduction at equilibrium, and "receiving" some of the perturbations resulting from nucleotide alterations. Disappearance of P-loop flexibility upon GTP heating despite mutual proximity suggests there is stronger allosteric GTP–P-loop–G domain coupling prior to hydrolysis than after it, when the P-loop (*cf.* D-NEMD scores in the purple-to-red range) only "feels" a moderate perturbation that it doesn't pass onwards to other K-Ras4B regions. As noted in the dedicated subsection, this is certainly supported by DF analysis: while GTP itself is not considered, DF scores for each of the P-loop residues 10–14 (Figure 3; bottom) remain low (white) with respect to most other G domain residues, with switches I and II the only major exceptions on top of the HVR. This is also clear from the DF video in

Supporting Information. As we have indicated when commenting Figure S2, the SPM similarly predicts a moderate allosteric coupling, with the P-loop's shortness sitting around the 0.3 threshold.

Speaking of the HVR itself, while its presence in Figure 7 is limited to its first few residues because of the noise associated with ATD (*vide infra*), representations in Figure 3–Figure 5 show that it too is approached very differently by our chosen languages. On the one hand, DF analysis confirms a notoriously high flexibility / poor allosteric coordination (*cf.* slightly higher DF scores in the rightmost part of the second row of Figure 7 and intense blue color in Figure 3), and of course exclusion from ATD is due to the same reason. In contrast, SPM predictions go in the opposite direction (Figure 4; 'light' S_{norm} values in the upper right of Figure 7) suggesting HVR can play an important allosteric role despite its flexibility. In principle, D-NEMD simulations reflect SPM, detecting very little average deviation in HVR upon hydrolysis (and this is a testament to their better ability to cancel out non-allosteric noise): however, in light of the discordant results of the other three approaches, we cannot conclusively tell whether this is due to a lack of any allosteric communication with the hydrolyzing substrate, or to dissipation of allosteric signals into the membrane (although the latter option seems less likely).

Two further salient elements that are recognized by the languages in an intriguingly different way are catalytic residues Lys16 and Gln61, which play different but equally important roles in the hydrolytic mechanism.[39] Prior to hydrolysis, Gln61 helps position the nucleophilic water to attack GTP:Pγ (and we should specify that this is far more likely in the presence of GAP; Figure S1b), whereas before and during hydrolysis, cationic Lys16 lowers the free energy barrier by abstracting negative charge from the transition state. S_{norm} values for Gln61 are easier to decipher (Figure 7): while all show some degree of predicted allosteric relevance, D-NEMD exceeds the others, showing a particularly strong allosteric perturbation resulting from hydrolysis (Figure 7;

third row; orange). This is entirely consistent with its dual role as a facilitator of nucleophilic attack (prior to hydrolysis), and (after hydrolysis) as a component of the perturbed switch II. In the case of Lys16, we assumed that any greater allosteric importance would be detected at least prior to hydrolysis, as the active site preorganizes to stabilize the transition state; however, in spite of the relatively high correlation with the G domain reported in the 2D **DF** matrix (Figure 3; bottom, Supplementary Video), three scores out of four only point to a moderate degree of allosteric importance (Figure 7; purple), and the SPM (Figure 7; top row; black band) shows a high degree of *de*coupling from the main allosteric path. This suggests that while crucial for catalysis, Lys16 did not evolve to become as prominent an allosteric regulator of it.[39]

Concluding with an overview of the discrepant secondary structure elements that aren't part of switches, we note first of all that sheet β1 is featured prominently in the SPM, and shows areas of strong coupling to the nucleotide binding site (ATD), but this is reflected neither in the 1D DF scores—as also confirmed by projections in Figure 3—nor by D-NEMD. The α1 helix, which besides Lys16 contains binding site residue Leu19 (mentioned in the SPM analysis), is found to be perturbed in D-NEMD simulations in several spots; ATD and 1D DF scores on the other hand only recognize a moderate allosteric value. Predictions for sheets β2 and β3 are influenced by their location between the strongly decoupled switch I and β2-β3 loop, for which consensus was universal across methods: there appears to be a slight allosteric stabilization for β2 but nothing more. While unstructured, we should also mention two loops that are not so important at equilibrium (black or purple in SPM and DF scores; Figure 7), but feel significant perturbation upon GTP heating or hydrolysis: these are the β5-α4 and β6-α5 loops at the guanine end of the nucleotide binding site; in fact, the former loop is one of the regions outside switches I and II that feels the greatest perturbation from hydrolysis *per* D-NEMD (*cf*. bright yellow band).

Finally, helix α5 merits a separate discussion, in accordance with the frequent clues to its centrality in K-Ras4B allostery that have emerged throughout the preceding analyses. The helix should ideally be mentioned at the opening of the previous subsection along with sheets β4 and β5 as a strong contributor to internal allostery at equilibrium that is unscathed by GTP perturbations: the reason why the helix is discussed here is merely because ATD scores are outliers, likely due to contamination from HVR noise. The remaining three approaches concur in confirming helix α5 as a fundamental allosteric player: average DF scores are the lowest (blackest) in the entire protein, signaling strong allosteric coupling; SPM shows high 'closeness' values (red bands) for crucial residues Ile163, Arg164, and Lys167; whereas D-NEMD shows moderate perturbations at similar positions along the helix, despite its distance from the active site.

The General Picture: The Membrane Talks to GTP, GTP Talks to the Switches. After piecing together the considerable amount of information extracted and commented thus far, there emerges a chronologically clear and articulate allosteric portrait of K-Ras4B (and one that is also experimentally consistent, as discussed in the next subsection). For the chosen GTP-active state of K-Ras4B[33] (*vide infra*), languages consistently describe a compact kinase in which allosteric signals appear to be far-reaching but rigorously compartmentalized: prior to hydrolysis, at equilibrium, a set of pathways 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.

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. Said binding site residues are first and foremost the P-loop and guanine-binding residues Ser145 and Ala146 on the β6-α5 loop.

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 if equilibration is considered). For reasons we discuss further below, 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[39] of some allosterically relevant 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 (*vide infra*).

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 (DNEMD; Figure 5) and of course mediate K-Ras4B inactivation, could not be clearer; while in proximity to the hydrolytic center, P-loop and Lys16 remain comparatively less perturbed by hydrolysis: we will comment in the next subsection on the intriguing links between hydrolysis and the β5–α4 loop, which is on the other side of the active site compared to the reaction center and known to experimentalists[38,54] (G4 motif, *vide infra*). We will also discuss implicit experimental consensus in the blander links to regions that are more distant, particularly

the *C*-terminus of helix α1, the β2 sheet outside switch I, and the *N*-terminus of helix α3. 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.

Dialoguing with Experiment: Switches, Pockets, Interfaces, Reactivity. Perhaps the most significant final validation of the articulate allosteric information emerging from our babel of methods is the general degree of consistency with the most recent experimental (and indeed computational) allosteric narrative of K-Ras4B, which we have thus far only marginally referred to. We will specifically focus on links to areas where conformational change is known to occur, structural flexibility or rigidity, pockets, interfaces, and reactivity.

Prior to any discussion, it is first essential we better characterize the conformational state of our chosen starting structure (PDB ID: 6vjj;[33] Figure 1) in relation to the several[36,38,51] that the kinase is known to adopt even when GTP-bound and active. Specifically, switch I is in the conformational effector-binding state typically reported as "state 2",[36,38] defined by the coordination of its Thr35 to Mg^{2+},[42] which is also electronically important to GTP reactivity.[39] The conformation of switch II features an ordered and well-resolved α2 helix, with the Gln61 projected towards the active site,[33] in an *active* conformation reported as "R state" that is opposed to an *inactive*, disordered "T state";[38] we note that *per* the R state, helix α3 and α3-β5 loop, which form an allosteric switch,[38] should be closer to helix α4, and the kink in helix α3 straightened: it is clear from Figure 1 that this is not the case with 6vjj.[33] Nonetheless, we still consider 6vjj to be closer to the R state.

The most evident structural consistency with previous studies is the substantial flexibility of switches I and II at equilibrium, with mild (but certain) allosteric links to the G domain, as well as

their sensitivity to hydrolysis, which denotes incipient conformational change. This is entirely in line with the reported existence of several conformational states even in the GTP bound form,[35,36,38,42,51] and in line with deactivation being linked to GTP hydrolysis, as active switch conformations become inaccessible. More specifically, at equilibrium, while switches I and II are (expectedly) already very mobile, they retain mild allosteric coupling (*cf*. Figure 3; bottom) through Mg^{2+}-bound Thr35 in the former case and Thr58/Ala59/[Gln61] in the latter case: this is what allows them to form functional interfaces, as we shall see. Once hydrolysis occurs (*vide infra*), departure of the cleaved phosphate and ensuing loss of Mg^{2+} coordination severs even these residual allosteric links on the switches, and it is in all likelihood this hydrolysis-related event that definitively hampers their ability to bind to effectors.

Another aspect nicely consistent with experiment, naturally, is the role detected for HVR: while flexibility is evident from our simulations and in line with its irresolution in most of the available crystal structures,[36] the shortest path map from equilibrium MD (Figure 4) (but not DF analysis) confirms its allosteric dialogue with the G domain, in agreement with experimental reports.[35] In terms of reported[43,44] HVR-switch (and G-domain—membrane) contacts, analysis of our simulations reveals that centers of mass of remain distant in all our replicas, mostly in the 35-40 Å range. Nonetheless, throughout our MD simulations we do observe a high degree of K-Ras4B mobility across the surface of the lipid bilayer, with the kinase frequently changing simulation cells: this fits in with the established role[35] of K-Ras4B in recruiting effector monomers to the cellular membrane and bringing them together so that they can dimerize.

While the P-loop is experimentally renowned[36] to be the quintessential site for oncogenic (hyperactivating) mutations, all of our chosen methods invariably classify it as of moderate

allosteric importance: this suggests that mutations on this loop do not (solely) act through internal allosteric disruption of hydrolysis, but through other means of stabilization proposed below.

Still, barring the switches, the HVR, and some loops, all allosteric languages describe the G domain as an entity that is fairly compact allosterically; this too is reported experimentally[40,42] and, incidentally, it is why, contrarily to what is customary in the literature, we have not felt it necessary to refer to the effector and allosteric lobes into which the G domain is conventionally partitioned.[35,36,40-42] Going further in depth, the conspicuous study by Weng *et al*. mentioned in the introduction,[40] thanks to point mutations systematically introduced across the entire K-Ras4B structure, specifically speaks of anisotropic allosteric coupling across the central beta sheets: this is in exact agreement with the recognition of the central allosteric hub by all four of our chosen methods, in some form, which indeed comprises the β4–β6 sheets. More encouragingly still, when the authors[40] go on to dissect which positions are most allosterically affected by mutations, they provide an explicit list of 8 'novel' allosteric points at various locations in the G domain (on top of 10 in the vicinity of the binding pocket); strikingly, 6 out of 8 of these novel allosteric sites are actually comprised in the SPM (Figure 4): Val7 (β1), Thr58, Ala59 (switch II), and—most importantly—Pro110 (α3-β5), Phe141 (β6), and Ile163 (α5), which precisely emerge as some of the "closest" (most communicating) residues in the shortest path map as part of the allosteric hub.

Mention of the α3-β5 loop and helix α5 in a number of experimental contexts[33,38,40] prompts us to shed more light on the allostery detected by our methods in these areas. Indeed, the reader will recall that the *C*-terminal portion of the α3-β5 loop, which we have just mentioned as part of the allosteric switch that straightens helix α3 and brings switch II closer to state R,[38] was the major area of universal consensus across languages appearing in Figure 7 along with sheet β5, with a high level of allosteric communication and low sensitivity to hydrolysis. Contrast with the

immediately preceding *N*-terminal portion and helix α3 itself could not be more evident (Figure 7): this portion has low allosteric communication at equilibrium (lighter DF, darker SPM), but is clearly more sensitive to hydrolysis (lighter D-NEMD). These very distinct behaviors are definitely compatible with the experimentally recognized nature of the α3/α3-β5 switch: the *C*-terminal portion holds the switch in place in the active state; however, once the *N*-terminal portion and helix α3 itself feel the effects of GTP hydrolysis, that part of the loop experiences new allosteric perturbations that will alter its conformation towards inactive state T.

Elsewhere on the allosteric hub that transduces signals from the HVR, our own recognition of the allosteric importance of helix α5—which features Ile163 that Weng *et al*.'s mutagenesis study also recognizes as allosterically prominent—is again in line with experiment. To recall, not only does the helix have several other sites appearing as "close" in the SPM, but it is also recognized as allosterically coupled with the remainder of the G domain (Figure 3; whiter **DF** matrix) and unperturbed by hydrolysis (Figure 7; darker D-NEMD). This is all the more interesting if one notes that helix α5 also plays a well-established role[33] in the interaction between K-Ras4B and part of RAF1, which is one of the most prominent effectors: allosteric rigidity at equilibrium is thus compatible with an interface that needs to remain well defined in the active state (PDB ID: 6xi7;[33] especially in the presence of more 'volatile' interfaces such as that with switch I). The important role for this helix in maintaining the GTP-active state is also confirmed in studies of pan-mutant inhibitor BI-2865[54] (*cf*. *Introduction*), wherein certain evidently 'allostery-reactivating' mutations in helix α5 are seen to diminish its effects despite the inhibitor—which favors the inactive GDP state—binding far from the helix.

Based on mutagenesis, the same study[54] also mentions specific allosteric cross-talk between residues Pro121/Ser122 on the β5-α4 loop and Asn85 on the β4-α3 loop as being determinant to

preserve the inactive state: examining our simulations, albeit conducted in the *active* state, 2D DF scores (Figure 3) certainly do not disagree with this, however, this is principally due to the generalized G domain compactness. It is more interesting, instead, to shift our attention on two other areas, namely, the conserved[38,54] G4 and G5 motifs (respectively: NK*x*D, residues 116-119, β5-α4 loop; and E*x*SAK, residues 143-147, β6 sheet/β6-α5 loop). Similarly to helix α5, in these spots too, certain point mutations are observed[54] to hamper BI-2865 activity, likely by favoring return to allosteric conditions similar to the GTP-active state. While G4 and G5 both feel an allosteric impulse from hydrolysis despite being on the opposite end of the binding site with respect to the cleaving γ-phosphate (Figure 5), we note that their conformation in the inactivated structure (PDB ID: 8azv)[54] is, like that of helix α5, unaffected. Any correspondence between our simulations and experiment must, rather, be sought in the allosteric data at equilibrium: G4 does not appear to be allosterically relevant at equilibrium, even though it must be said that its role is prevalently chemical rather than allosteric, since is role is to aid in the recognition of GTP over ATP;[38] on the other hand, like parts of helix α5, all G5 residues invariably appear in the SPM (branch II Figure 4; bottom) denoting a pronounced allosteric relevance in the GTP-active state at equilibrium that the GDP-inactive state evidently does not possess.

More in general, we should extend our analysis to the four allosteric pockets previously recognized by Grant *et al*. through simulations[41] and later validated experimentally,[40] in order to evaluate how well our own recipe is able to recognize these pockets, and thus assess its potential in predicting allosteric pockets in new systems. (In the designation used by both studies,[40,41] inhibitors Sotorasib and BI-2865 both bind to pocket II). While of course interference with all four pockets ultimately ushers in diminished K-Ras4B activity, our own assessment reveals that pocket

II,[41] pocket IV,[41] and pockets I/III[41] fall, respectively, into three distinct categories of which the latter is conceptually distinct from the former two.

Firstly, Sotorasib/BI-2865 binding pocket II should be envisaged as a stabilizer of the GDP-inactive state, which was indeed considered alongside the GTP state in Grant *et al*.'s simulations:[41] taking advantage of the conspicuous conformational shift in switch II after GTP hydrolysis, the pocket occupies the volume created between inactivated switch II and allosteric helix α3. The shallow pocket IV is found[41] through blind docking on GTP-active K-Ras4B, and occupies the volume between switch I, helix α1, and the rest of sheet β2: while this could lead to believe that the pocket stabilizes the harmful active state, it actually occupies an area that (*cf., e.g*. PDB ID: 6xi7)[33] is crucial for the interaction with effector RAF1, meaning that the pocket's occupation would effectively limit harmfulness of the GTP-active state. Strictly for the purposes of our discussion, what pockets II and IV have in common is that they require some prior degree of knowledge either about other biologically relevant conformational states, or about interfaces: therefore, their identification is not possible from our simulations alone and in fact, they consistently involve areas of low allosteric relevance at equilibrium.

On the other hand, the radically different nature of pockets I/III means that they are fully identified by our chosen approach and simulations: as clearly inferable especially from SPM (Figure 4) and D-NEMD data (Figure 5), they occupy regions of the 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 β1, β3, and the *C*-terminus of helix α5; and pocket III spans none other than the main α3–β5/β5/α5 allosteric hub. Interference with these pockets is thus bound to bring maximum disruption to the active state, and experimental data certainly suggests so too.[40] This correspondence with experiment bodes well for novel systems

whose allosteric properties are not as pervasively known: in such cases, one could for example compare and contrast the information provided by the SPM and D-NEMD to predict binding sites that could disrupt allostery if interfered with; in systems that are allosterically less compact, and wherein areas of allosteric coupling and decoupling are more easily distinguishable, DF analysis has certainly proven to be a useful guide too.[57]

Turning to the impact of allostery on K-Ras4B interfaces, we have already seen that helix α5 forms an important *rigid* part of the interface with effector RAF1 (PDB ID: 6xi7),[33] alongside switch I, helix α1, and the remainder of sheet β2 / loop α1-β2, which on the other hand D-NEMD simulations (Figure 5) show to be perturbed *post*-hydrolysis (and that we have seen to be part of pocket IV).[41] The presence of areas in the K-Ras4B—RAF1 interface that are clearly disrupted by GTP hydrolysis is experimentally logical, and provides the molecular explanation for disabled RAF1 binding by the GDP-inactive state. In fact, in five other Ras-effector interfaces dissected by Weng *et al.*,[40] the six residues that are in common to all and are most impacted by mutation (25, 33, 36, 38-40) are located in α1, α1-β2, and β2, signaling that the perturbation of this wider area during hydrolysis (and not just switch I) has actual biological repercussions; indeed, inactive "state 1" of switch I also entails open β2 and β3 sheets.[36] The ability of D-NEMD to capture interface perturbations in general[17] is also reconfirmed through these simulations.

Even the interface with GAP (Figure S1b; *cf.* PDB ID: 1wq1),[66] which is not an effector but oversees hydrolysis and therefore requires preference for the GTP-active state, spans this perturbed α1/α1-β2/β2 area that includes switch I. In addition (and this of course is in common with other effectors), GAP also interfaces with the equally perturbed switch II and loop β3-α2, and other marginally perturbed (Figure 5) but crucial areas including, notably, the P-loop (wherein mutations could therefore have a sterically disruptive effect for this deactivating interaction).

We conclude our comparison with previous studies by examining potential links between allostery, mutations, and reactivity suggested by our own simulations. A good source of guidance is the QM/MM investigation by Calixto *et al.*[39] on the multifaceted role of GAP (Figure S1b) in facilitating GTP hydrolysis. As we have amply discussed, switches I and II morph to recruit regulatory proteins such as RAF1 and GAP, but remain allosterically controlled by Mg^{2+}-bound Thr35 on switch I and (ordered) Thr58/Ala59 on switch II. Looking at reactivity, it is moreover evident how important these residues are in exerting control on hydrolysis itself: a Mg^{2+} that is immobilized by Thr35 is essential to coordinate GTP and stabilize the transition state,[39] just as the widely discussed Gln61, downstream of Thr58/Ala59, is essential to immobilize the attacking nucleophile[39] (*cf. Introduction*). It is also similarly clear why, once inorganic phosphate is released after hydrolysis, allosteric links to these switches are severed.

As validated catalytic actors,[39] Thr35 and Gln61, have, together with Lys16, been abundantly considered in our discussion, in anticipation of the fact that they would also be the main mediators of allosteric control on hydrolysis; however, the comparatively more modest allosteric value predicted for these residues leads us to speculate that there should be additional control mechanisms. An intriguing hypothesis as to where this supplemental control could come from emerges if one considers favorable and unfavorable electrostatic contributions to hydrolysis[39] by less obvious residues, including SPM node Ala59, and borderline SPM node Gly13 on the P-loop (which has so far only been mentioned in the context of its proneness to mutate). Mutations on the P-loop could thus not only disrupt allosteric control on hydrolysis, but also, more simply, disrupt reaction electrostatics, or (*e.g.*, G12D) sequester the catalytic arginine administered by GAP. Proving these hypotheses, however, would require additional simulations and is beyond the scope of this work.

The Importance of Learning many (Allosteric) Languages to Decipher Unbiased MD. One final point to review are the reasons and implications of our methodological choices. With so many different *in silico* "allosteric languages" available across the scientific community, some now for many years (*cf. Introduction*), it was clearly beyond the scope of this work to exhaustively compare-and-contrast as many methods as possible. Rather, we have opted to limit ourselves to four methods—quite different in scope and theoretical origins—and 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 *in silico* allostery detection methods can be in capturing them adequately.

We should stress in the clearest possible way that our choice of methods does not imply an indication of preference or support for any one of them; in fact, while compelled to seek methods that were as heterogeneous as possible (*i.e.*, the linguistic nuance that we have discussed at length), all authors of this work have made it a point to employ approaches that were beyond their immediate areas of expertise. Much like when learning any new language, this has an inevitable temporal downside as one familiarizes with the subtleties, advantages, and pitfalls of a particular approach; nonetheless, it is still an investment worth making, and we heartily recommend it to authors wishing to undertake a new allosteric study. Downsides to monitor closely while seeking to maximize heterogeneity include, for example, excessive computational cost, persistence of noise in the allosteric data, user-unfriendliness, and poor automatability or portability.

To briefly exemplify some of the considerations specifically applying to our own choice of methods, SPM and DF clearly have the greatest advantage in terms of computational cost since, as we recounted earlier, they can extract allosteric information from just one series of unbiased

MD simulations (D-NEMD and ATD require additional integrative simulations). "Equilibrium" languages SPM and DF both have their own unique strengths and weaknesses. For example, DF allows a unique two-dimensional breakdown of allosteric cross-talk signals at equilibrium (the **DF** score matrix; Figure 3, bottom), which as we have seen is useful, since not all allosterically active regions dialogue with each other with the same intensity. On the other hand, 'flattening' the **DF** matrix leads to an average 1D score that can be projected onto a structure under examination (*e.g.* Figure 3, top/middle), but loses a considerable amount of spatial information, *i.e.*, one only retains information on tendentially rigid *vs.* tendentially uncoordinated residues. The SPM approach provides an elegant way to circumvent this problem: while still starting from the multidimensional information about how each residue talks to every other residue, the allosteric maps that are produced (*e.g.*, Figure 4) conveniently condense the information into a linear 'shortest path', which obviously not only retains spatial information about the "flow" of allosteric signals, but also quantitative information about their changing intensity.

One *caveat* of SPM, however, is that being based on average distances and covariances (*cf. Computational Methods*) it requires that they be measured after structural alignment to a sufficiently representative reference structure: this needs to be derived carefully, as the choice of clustering and alignment can somewhat impact results, particularly in path sections close to the 0.3 'shortness' threshold: appearance of the P-loop in Figure S2c but not Figure S2a is a good example of this. The *caveat* doesn't apply to DF analysis, which only monitors intra-frame Cα–Cα distances regardless of orientation, and therefore only requires that a molecule in a particular frame is whole.

While perturbative methods D-NEMD and ATD require extra simulations and therefore come at an increased computational cost, it is a price worth paying given the level of allosteric insight

that they are able to provide: with the notable advantage of not requiring to model a conformational change, both these languages can offer an accurate indication of what happens allosterically upon a specific perturbation, when only a part of the equilibrium pathways will be exalted over others. Though it may be argued that the perturbation introduced by these methods is a form of bias, this is not strictly true, insofar as they do not accelerate a particular conformational transition but, still retaining equilibrium MD simulations as their standard, they merely try and detect which areas of the system vary the most once the perturbation is introduced. Conveniently, the perturbation does not have to make strict physicochemical sense, as is the case with both the supercooling and decoupled heating in ATD and the instantaneous GTP hydrolysis in D-NEMD.

A particularly elegant aspect of D-NEMD (Figure 2) is that the non-equilibrium simulations are initiated instantaneously from equilibrium ones, with the same set of atomic velocities, and deviation is monitored systematically at synchronous intervals. This essentially suppresses any noise that may arise from inherently flexible regions, because flexible regions that aren't affected by the perturbation under examination will always deviate by identical amounts both at equilibrium and at non-equilibrium; the only appreciable source of deviation between the two situations is thus bound to be ascribable to the perturbation itself.

ATD, on the other hand, is less capable of eliminating noise, since deviations in each window are monitored from a fixed time reference (*i.e.*, the start of heating for that window): this means that it is impossible to entirely separate deviations that simply happen because of flexibility, and those that happen because they have received an allosteric impulse from the perturbation, and it is the reason why the HVR had to be excluded from the S_{norm} plots in Figure 7. The attractivity of ATD lies in the fact that it can act as a sort of 2D "allosteric NMR", selectively overstimulating

the contribution of one particular residue, sets of residues, or as in our case, of course, the nucleotide itself.

Finally, an inevitable downside common to all four methods is that they report "average" effects. In other words, while originating form a single structure, unbiased MD will still lead the system under study to explore a variety of conformations: it may well be that certain allosteric signals in a particular set of conformations are more (or less) pronounced than in others, but it is worth remembering that this won't show up when looking at averages.

We should also obviously stress stress that plenty of other options would have been available for our investigations. While it is not possible to review them here in detail they should by all means be considered as alternatives by any reader wishing to plan their own investigation. A few examples taken from across the realms of equilibrium and perturbative allostery alike include Normal Mode Analysis,[90] Gaussian Network Models,[91] Leverage Coupling,[21] Perturbed Ensemble Analysis,[92] and Networks of Local Correlated Motions.[93] Additional methods and online resources are reviewed in more detail elsewhere.[5,21,23] It is also important to point out that findings from a chosen combination of languages could be further reinforced by forms of systematic coevolution analysis[2,26] and/or analysis of the most frequent pathogenic mutation sites.[94]

Returning to our own choice, and especially 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 and fairly portable to other systems—can work synergistically to provide a high level of allosteric detail about a particular system that they would be unable to provide separately. 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: indeed, the large amount of data originating from this work is, essentially, derived from just a single starting structure. Alternative approaches to study different systems should 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 timescales 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 kinase 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 K-Ras4B,[35,36,40] with its notoriously difficult pharmaceutical targetability, was deliberately chosen to prove the benefits of this combination of methods. It is an oncotarget characterized by pervasive allosteric effects: normally inactivated by GTP hydrolysis, which incapacitates the switches, it can be malignantly hyperactivated by a series of oncogenic mutations that block allosteric triggers of hydrolysis.

Languages chosen to interrogate the active state of K-Ras4B at equilibrium include: (i) distance fluctuation analysis,[14,16,57] which postulates that pairs of residues moving more concertedly than others—with distances remaining closer to the simulation average—should represent hotspots of allosteric change; and (ii) the shortest path map,[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) dynamic non-equilibrium MD[17,58-63] spawns 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 form equilibrium MD, averaged over all short simulations, reflects the degree of allosteric perturbation induced by hydrolysis. Finally, (iv) anisotropic thermal diffusion[64] 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 predicting high or low allosteric importance for certain areas of K-Ras4B, they provide intriguingly different answers for certain other areas: it is thanks to these linguistic nuances that 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, whencefrom they reach a paramount allosteric hub centered around the *C*-terminal half of (tendentially rigid) helix α5, sheet β5, and around loop α3–β5. From here, they branch out to encompass most remaining parts of the kinase, notably including—albeit to moderate degrees—residues on several sides of the GTP binding site. Amongst these, there 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 QM/MM reactivity studies,[39] 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 presence of the above allosteric links, switches themselves exhibit low allosteric coordination, in agreement with their interfacial plasticity. Upon hydrolysis, most of the kinase remains unaffected, except for the two switches (on which it is known to have an inactivating impact), and the α1/α1-β2/β2 area: all belong to crystallographically known interfaces.[33,66]

Consistency with experiment is recognizable in a number of other aspects, notably in the importance of the α5/β5/α3–β5 allosteric hub,[40] general allosteric compactness,[40,42] and retention of allosteric links in the hypervariable region and in both switches.[36] Our data also agree with experiment in terms of the allosteric relevance of helix α5 and the G5 domain;[54] and two of four pockets suggested by mutagenesis[40] 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, in confirmation of our initial argument, 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, furthermore,

decryption of such an articulate portrait was possible even if our simulations were all begun from a single kinase 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 to biocatalysts, regardless of the computational "languages" used to decipher it. We believe our work—which presents one of the first systematic comparisons of different allostery detection methods—should inspire biosimulation communities to choose, learn, and apply as many "allosteric languages" as possible, and with as much creativity as computational resources permit.

ASSOCIATED CONTENT

Supporting Information. The following files are available free of charge.

Depiction of the starting structure and the K-Ras4B—GAP complex; MD preproduction details; Alignment, clustering, and matrix derivation for SPM; SPM drawn on the full 5 μs metatrajectory; Further D-NEMD details and "reactive" pose counts. (PDF)

Equilibrium MD: starting coordinates, topology, input; ATD: input; D-NEMD: topology and input; SPM: alignment scripts, input matrices for the 4.9 and 5 μs metatrajectories, output. (ZIP)

D-NEMD: average Cα deviation from 0 to 50 ps, projected on 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)

AUTHOR INFORMATION

Corresponding Author

*To whom correspondence should be addressed: g.colombo@unipv.it, stefanoartin.serapian@unipv.it

Author Contributions

SAS and GC conceived the work and planned the experiments. SAS prepared the initial topology as received from previous simulations, supervised the work, drew Figures, and applied the D-NEMD method. MC set up and carried out equilibrium molecular dynamics simulations, thereafter conducting DF analysis and helping with the Figures. FM carried out molecular dynamics simulations with Anisotropic Thermal Diffusion. All authors have written and reviewed parts of the manuscript and agree with the contents of this study.

Funding Sources

Authors wish to thank AIRC (IG 27139) for funding.

Acknowledgements

SAS wishes to thank Prof. Alemayehu Gorfe (Univesity 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 forcefield parameters, and Prof. Aleksander Lyubartsev (Stockholm University) for assistance with the *Slipids* forcefield. Special thanks are also due to Dr. Ana Sofia Fernandes Oliveira (University of Bristol) for her precious advice on the D-NEMD method, and Dr. Sílvia Osuna Oliveras (University of Girona) for kindly providing her SPM code. All authors acknowledge support from the Italian *Ministero dell'Università e della Ricerca* (MUR) and the University of Pavia 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. DOI: 10.1126/science.286.5438.295.

(2) Granata, D.; Ponzoni, L.; Micheletti, C.; Carnevale, V. Patterns of coevolving amino acids unveil structural and dynamical domains. *Proceedings of the National Academy of Sciences* **2017**, *114* (50), E10612-E10621. DOI: 10.1073/pnas.1712021114.

(3) Orozco, M. A theoretical view of protein dynamics. *Chemical Society Reviews* **2014**, *43* (14), 5051-5066, 10.1039/C3CS60474H. DOI: 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. *Chemical Reviews* **2016**, *116* (11), 6391-6423. DOI: 10.1021/acs.chemrev.5b00623.

(5) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450* (7172), 964-972. DOI: 10.1038/nature06522.

(6) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508* (7496), 331-339. DOI: 10.1038/nature13001.

(7) Gunasekaran, K.; Ma, B.; Nussinov, R. Is allostery an intrinsic property of all dynamic proteins? *Proteins: Structure, Function, and Bioinformatics* **2004**, *57* (3), 433-443. DOI: https://doi.org/10.1002/prot.20232.

(8) Goodey, N. M.; Benkovic, S. J. Allosteric regulation and catalysis emerge via a common route. *Nature Chemical Biology* **2008**, *4* (8), 474-482. DOI: 10.1038/nchembio.98.

(9) Liu, J.; Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Computational Biology* **2016**, *12* (6), e1004966. DOI: 10.1371/journal.pcbi.1004966.

(10) Tsai, C.-J.; Nussinov, R. A Unified View of "How Allostery Works". *PLOS Computational Biology* **2014**, *10* (2), e1003394. DOI: 10.1371/journal.pcbi.1003394.

(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. *Nature Chemistry* **2021**, *13* (10), 1017-1022. DOI: 10.1038/s41557-021-00763-6.

(12) Romero-Rivera, A.; Garcia-Borràs, M.; Osuna, S. Role of Conformational Dynamics in the Evolution of Retro-Aldolase Activity. *ACS Catalysis* **2017**, *7* (12), 8524-8532. DOI: 10.1021/acscatal.7b02954.

(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 Catalysis* **2015**, *5* (2), 1172-1185. DOI: 10.1021/cs501704f.

(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. *Journal of Molecular Biology* **2023**, *435* (3), 167931. DOI: https://doi.org/10.1016/j.jmb.2022.167931.

(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. *Frontiers in Oncology* **2020**, *10* (1177), Review. DOI: 10.3389/fonc.2020.01177.

(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. *Journal of Molecular Biology* **2022**, *434* (17), 167468. DOI: https://doi.org/10.1016/j.jmb.2022.167468.
(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. *The European Physical Journal B* **2021**, *94* (7), 144. DOI: 10.1140/epjb/s10051-021-00157-0.
(18) Monod, J.; Wyman, J.; Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology* **1965**, *12* (1), 88-118. DOI: https://doi.org/10.1016/S0022-2836(65)80285-6.
(19) Cooper, A.; Dryden, D. T. F. Allostery without conformational change. *European Biophysics Journal* **1984**, *11* (2), 103-109. DOI: 10.1007/BF00276625.
(20) Cui, Q.; Karplus, M. Allostery and cooperativity revisited. *Protein Science* **2008**, *17* (8), 1295-1307. DOI: https://doi.org/10.1110/ps.03259908.
(21) Guarnera, E.; Berezovsky, I. N. Allosteric sites: remote control in regulation of protein activity. *Current Opinion in Structural Biology* **2016**, *37*, 1-8. DOI: https://doi.org/10.1016/j.sbi.2015.10.004.
(22) Nussinov, R.; Tsai, C.-J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153* (2), 293-305. DOI: 10.1016/j.cell.2013.03.034 (acccessed 2022/12/16).
(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. *Chemical Reviews* **2016**, *116* (11), 6516-6551. DOI: 10.1021/acs.chemrev.5b00562.
(24) Gianni, S.; Jemth, P. Allostery Frustrates the Experimentalist. *Journal of Molecular Biology* **2023**, *435* (4), 167934. DOI: https://doi.org/10.1016/j.jmb.2022.167934.
(25) Nussinov, R.; Tsai, C.-J. Unraveling structural mechanisms of allosteric drug action. *Trends in Pharmacological Sciences* **2014**, *35* (5), 256-264. DOI: https://doi.org/10.1016/j.tips.2014.03.006.
(26) Crean, R. M.; Gardner, J. M.; Kamerlin, S. C. L. Harnessing Conformational Plasticity to Generate Designer Enzymes. *Journal of the American Chemical Society* **2020**, *142* (26), 11324-11342. DOI: 10.1021/jacs.0c04924.
(27) 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. *Chemical Communications* **2018**, *54* (50), 6622-6634, 10.1039/C8CC02426J. DOI: 10.1039/C8CC02426J.
(28) Serapian, S. A.; Colombo, G. Designing Molecular Spanners to Throw in the Protein Networks. *Chemistry – A European Journal* **2020**, *26* (21), 4656-4670. DOI: 10.1002/chem.201904523.
(29) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. Fine-Tuning Multiprotein Complexes Using Small Molecules. *ACS Chemical Biology* **2012**, *7* (8), 1311-1320. DOI: 10.1021/cb300255p.
(30) 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. *Journal of Molecular Biology* **2006**, *356* (3), 746-758. DOI: https://doi.org/10.1016/j.jmb.2005.11.085.

(31) Osuna, S. The challenge of predicting distal active site mutations in computational enzyme design. *WIREs Computational Molecular Science* **2021**, *11* (3), e1502. DOI: https://doi.org/10.1002/wcms.1502.
(32) 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. *Nature Reviews Methods Primers* **2021**, *1* (1), 46. DOI: 10.1038/s43586-021-00044-z.
(33) 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. *Nature Communications* **2021**, *12* (1), 1176. DOI: 10.1038/s41467-021-21422-x.
(34) 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. *Biophysical Journal* **2016**, *110* (5), 1125-1138. DOI: https://doi.org/10.1016/j.bpj.2016.01.019.
(35) Nussinov, R.; Jang, H.; Tsai, C.-J.; Liao, T.-J.; Li, S.; Fushman, D.; Zhang, J. Intrinsic protein disorder in oncogenic KRAS signaling. *Cellular and Molecular Life Sciences* **2017**, *74* (17), 3245-3261. DOI: 10.1007/s00018-017-2564-3.
(36) Pantsar, T. The current understanding of KRAS protein structure and dynamics. *Computational and Structural Biotechnology Journal* **2020**, *18*, 189-198. DOI: https://doi.org/10.1016/j.csbj.2019.12.004.
(37) Janosi, L.; Gorfe, A. A. Segregation of Negatively Charged Phospholipids by the Polycationic and Farnesylated Membrane Anchor of Kras. *Biophysical Journal* **2010**, *99* (11), 3666-3674. DOI: https://doi.org/10.1016/j.bpj.2010.10.031.
(38) 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. *Journal of Biological Chemistry* **2017**, *292* (31), 12981-12993. DOI: https://doi.org/10.1074/jbc.M117.778886.
(39) 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. *Journal of the American Chemical Society* **2019**, *141* (27), 10684-10701. DOI: 10.1021/jacs.9b03193.
(40) Weng, C.; Faure, A. J.; Lehner, B. The energetic and allosteric landscape for KRAS inhibition. *bioRxiv* **2022**, 2022.2012.2006.519122. DOI: 10.1101/2022.12.06.519122.
(41) 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), e25711. DOI: 10.1371/journal.pone.0025711.
(42) Chao, F.-A.; Dharmaiah, S.; Taylor, T.; Messing, S.; Gillette, W.; Esposito, D.; Nissley, D. V.; McCormick, F.; Byrd, R. A.; Simanshu, D. K.; et al. Insights into the Cross Talk between Effector and Allosteric Lobes of KRAS from Methyl Conformational Dynamics. *Journal of the American Chemical Society* **2022**, *144* (9), 4196-4205. DOI: 10.1021/jacs.2c00007.
(43) Chavan, Tanmay S.; Jang, H.; Khavrutskii, L.; Abraham, Sherwin J.; Banerjee, A.; Freed, Benjamin C.; Johannessen, L.; Tarasov, Sergey G.; Gaponenko, V.; Nussinov, R.; et al. High-Affinity Interaction of the K-Ras4B Hypervariable Region with the Ras Active Site. *Biophysical Journal* **2015**, *109* (12), 2602-2613. DOI: https://doi.org/10.1016/j.bpj.2015.09.034.
(44) 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. *The FASEB Journal* **2016**, *30* (4), 1643-1655. DOI: https://doi.org/10.1096/fj.15-279091.

(45) Kapoor, A.; Travesset, A. Differential dynamics of RAS isoforms in GDP- and GTP-bound states. *Proteins: Structure, Function, and Bioinformatics* **2015**, *83* (6), 1091-1106. DOI: https://doi.org/10.1002/prot.24805.
(46) López, C. A.; Agarwal, A.; Van, Q. N.; Stephen, A. G.; Gnanakaran, S. Unveiling the Dynamics of KRAS4b on Lipid Model Membranes. *The Journal of Membrane Biology* **2021**, *254* (2), 201-216. DOI: 10.1007/s00232-021-00176-z.
(47) 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. *Journal of Biological Chemistry* **2015**, *290* (48), 28887-28900. DOI: https://doi.org/10.1074/jbc.M115.664755.
(48) 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 Computational Biology* **2018**, *14* (9), e1006458. DOI: 10.1371/journal.pcbi.1006458.
(49) Prakash, P.; Gorfe, A. A. Lessons from computer simulations of Ras proteins in solution and in membrane. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830* (11), 5211-5218. DOI: https://doi.org/10.1016/j.bbagen.2013.07.024.
(50) Prakash, P.; Gorfe, A. A. Probing the Conformational and Energy Landscapes of KRAS Membrane Orientation. *The Journal of Physical Chemistry B* **2019**, *123* (41), 8644-8652. DOI: 10.1021/acs.jpcb.9b05796.
(51) 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. *Physical Chemistry Chemical Physics* **2022**, *24* (13), 7653-7665, 10.1039/D2CP00274D. DOI: 10.1039/D2CP00274D.
(52) Zhang, M.; Jang, H.; Nussinov, R. The mechanism of PI3Kα activation at the atomic level. *Chemical Science* **2019**, *10* (12), 3671-3680, 10.1039/C8SC04498H. DOI: 10.1039/C8SC04498H.
(53) 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. DOI: 10.1038/s41586-019-1694-1.
(54) 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**. DOI: 10.1038/s41586-023-06123-3.
(55) Petrović, D.; Kamerlin, S. C. L. Molecular modeling of conformational dynamics and its role in enzyme evolution. *Current Opinion in Structural Biology* **2018**, *52*, 50-57. DOI: https://doi.org/10.1016/j.sbi.2018.08.004.
(56) 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. *International Journal of Molecular Sciences* **2023**, *24* (9), 7747.
(57) 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 Computational Biology* **2012**, *8* (3), e1002433. DOI: 10.1371/journal.pcbi.1002433.
(58) Ciccotti, G.; Jacucci, G. Direct Computation of Dynamical Response by Molecular Dynamics: The Mobility of a Charged Lennard-Jones Particle. *Physical Review Letters* **1975**, *35* (12), 789-792. DOI: 10.1103/PhysRevLett.35.789.

(59) 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: Structure, Function, and Bioinformatics* **2020**, *88* (1), 113-126. DOI: https://doi.org/10.1002/prot.25776.
(60) 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 Science* **2011**, *20* (7), 1220-1230. DOI: https://doi.org/10.1002/pro.650.
(61) 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*, e66567. DOI: 10.7554/eLife.66567.
(62) 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 α4β2 Nicotinic Receptor: Insights from Equilibrium and Nonequilibrium Simulations. *Structure* **2019**, *27* (7), 1171-1183.e1173. DOI: 10.1016/j.str.2019.04.008 (acccessed 2021/07/21).
(63) 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**. DOI: 10.1021/jacsau.3c00185.
(64) Ota, N.; Agard, D. A. Intramolecular Signaling Pathways Revealed by Modeling Anisotropic Thermal Diffusion. *Journal of Molecular Biology* **2005**, *351* (2), 345-354. DOI: https://doi.org/10.1016/j.jmb.2005.05.043.
(65) Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.
(66) 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. DOI: doi:10.1126/science.277.5324.333.
(67) 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. *Journal of Computational Chemistry* **2005**, *26* (16), 1668-1688. DOI: doi:10.1002/jcc.20290.
(68) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation* **2011**, *7* (2), 525-537. DOI: 10.1021/ct100578z.
(69) 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. DOI: https://doi.org/10.1016/j.softx.2015.06.001.
(70) 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: Structure, Function, and Bioinformatics* **2006**, *65* (3), 712-725. DOI: https://doi.org/10.1002/prot.21123.
(71) 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: Structure, Function, and Bioinformatics* **2010**, *78* (8), 1950-1958. DOI: https://doi.org/10.1002/prot.22711.

(72) 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. *Journal of Chemical Theory and Computation* **2013**, *9* (12), 5653-5674. DOI: 10.1021/ct400556v.
(73) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *The Journal of Physical Chemistry B* **2008**, *112* (30), 9020-9041. DOI: 10.1021/jp8001614.
(74) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion–Water Coordination and Exchange in Biomolecular Simulations. *Journal of Chemical Theory and Computation* **2012**, *8* (4), 1493-1502. DOI: 10.1021/ct3000734.
(75) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of polyphosphate parameters for use with the AMBER force field. *Journal of Computational Chemistry* **2003**, *24* (9), 1016-1025. DOI: 10.1002/jcc.10262.
(76) Kashefolgheta, S.; Vila Verde, A. Developing force fields when experimental data is sparse: AMBER/GAFF-compatible parameters for inorganic and alkyl oxoanions. *Physical Chemistry Chemical Physics* **2017**, *19* (31), 20593-20607, 10.1039/C7CP02557B. DOI: 10.1039/C7CP02557B.
(77) Jämbeck, J. P. M.; Lyubartsev, A. P. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *Journal of Chemical Theory and Computation* **2012**, *8* (8), 2938-2948. DOI: 10.1021/ct300342n.
(78) Jämbeck, J. P. M.; Lyubartsev, A. P. Another Piece of the Membrane Puzzle: Extending Slipids Further. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 774-784. DOI: 10.1021/ct300777p.
(79) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79* (2), 926-935. DOI: 10.1063/1.445869.
(80) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012**, *5* (1), 367. DOI: 10.1186/1756-0500-5-367.
(81) Hockney, R. W.; Goel, S. P.; Eastwood, J. W. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics* **1974**, *14* (2), 148-158. DOI: https://doi.org/10.1016/0021-9991(74)90010-2.
(82) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126* (1), 014101. DOI: 10.1063/1.2408420.
(83) 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. *The Journal of Chemical Physics* **1984**, *81* (8), 3684-3690. DOI: 10.1063/1.448118.
(84) Darden, T.; York, D. M.; Pedersen, L. G. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98* (12), 10089-10092. DOI: 10.1063/1.464397.
(85) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18* (12), 1463-1472. DOI: https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.
(86) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **1992**, *13* (8), 952-962. DOI: doi:10.1002/jcc.540130805.
(87) https://github.com/colombolab/Distance-Fluctuation-DF-Analysis (accessed.

(88) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, *9* (7), 3084-3095. DOI: 10.1021/ct400341p.
(89) 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.; et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25* (11), 1422-1423. DOI: 10.1093/bioinformatics/btp163.
(90) Bahar, I.; Lezon, T. R.; Bakan, A.; Shrivastava, I. H. Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins. *Chemical Reviews* **2010**, *110* (3), 1463-1497. DOI: 10.1021/cr900095e.
(91) Li, H.; Chang, Y.-Y.; Yang, L.-W.; Bahar, I. iGNM 2.0: the Gaussian network model database for biomolecular structural dynamics. *Nucleic Acids Research* **2015**, *44* (D1), D415-D422. DOI: 10.1093/nar/gkv1236 (acccessed 7/13/2023).
(92) Yao, X.-Q.; Hamelberg, D. Detecting Functional Dynamics in Proteins with Comparative Perturbed-Ensembles Analysis. *Accounts of Chemical Research* **2019**, *52* (12), 3455-3464. DOI: 10.1021/acs.accounts.9b00485.
(93) Pandini, A.; Fornili, A.; Fraternali, F.; Kleinjung, J. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. *The FASEB Journal* **2012**, *26* (2), 868-881. DOI: https://doi.org/10.1096/fj.11-190868.
(94) 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 Research* **2014**, *43* (D1), D805-D811. DOI: 10.1093/nar/gku1075 (acccessed 7/13/2023).

SYNOPSIS AND TOC.

Do you speak Allosteric? Starting from unbiased MD simulations of the well-known oncotarget K-Ras4B, we assess performance of four different approaches in detecting its allosteric behavior. Each approach brings its own unique insights, and shows consistency with experiment.