



OPEN

KRAS(G12C)–AMG 510 interaction dynamics revealed by all-atom molecular dynamics simulations

Tatu Pantsar

The first KRAS(G12C) targeting inhibitor in clinical development, AMG 510, has shown promising antitumor activity in clinical trials. On the molecular level, however, the interaction dynamics of this covalently bound drug–protein complex has been undetermined. Here, we disclose the interaction dynamics of the KRAS(G12C)–AMG 510 complex by long timescale all-atom molecular dynamics (MD) simulations (total of 75 μ s). Moreover, we investigated the influence of the recently reported post-translational modification (PTM) of KRAS' N-terminus, removal of initiator methionine (iMet1) with acetylation of Thr2, to this complex. Our results demonstrate that AMG 510 does not entrap KRAS into a single conformation, as one would expect based on the crystal structure, but rather into an ensemble of conformations. AMG 510 binding is extremely stable regardless of highly dynamic interface of KRAS' switches. Overall, KRAS(G12C)–AMG 510 complex partially mimic the native dynamics of GDP bound KRAS; however, AMG 510 stabilizes the α 3-helix region. N-terminally modified KRAS displays similar interaction dynamics with AMG 510 as when Met1 is present, but this PTM appears to stabilize β 2– β 3-loop. These results provide novel conformational insights on the molecular level to KRAS(G12C)–AMG 510 interactions and dynamics, providing new perspectives to RAS related drug discovery.

KRAS is a driver oncogene that is observed particularly in pancreatic, colorectal and lung cancers¹. Most often, KRAS becomes oncogenic by a missense mutation in codon 12 which is coding glycine (G12). From the KRAS G12 missense mutants, G12D and G12V are the most frequent, followed by G12C at the third place². In lung cancers, however, smoking-associated G12C mutation is dominating^{3,4}. Overall, there is a high demand for mutant KRAS targeted therapies^{5,6}.

As a noncatalytic cysteine, G12C is a suitable target for pharmacological intervention via a covalent inhibition of a small molecule drug⁷, providing a potential strategy to target this oncogenic KRAS mutant protein. Since discovery of an allosteric pocket beneath switch-II⁸, determined efforts have been made in the development of G12C targeting covalent inhibitors. Few years later from this initial discovery, a cell-active G12C targeting covalent inhibitor was disclosed⁹, and finally in vivo activity was demonstrated with ARS-1620¹⁰. For more comprehensive overview of covalently reacting G12C targeting small molecules, the reader is recommended a recent review by Goody et al.¹¹.

Currently, G12C inhibitors have reached the clinical trials. The first covalent drug in the clinical development, AMG 510 (Fig. 1A), demonstrated promising antitumor activity¹². Another G12C targeting inhibitor, for which clinical data has been reported, is MRTX849 from Mirati Therapeutics¹³. Overall, four KRAS(G12C) targeting inhibitors are currently in the clinical trials. In addition to AMG 510, which is in Phase 1/2 as monotherapy (NCT03600883) and in Phase 1 as combination therapy (NCT04185883¹⁴) and MRTX849 in Phase 1/2 (NCT03785249), Johnson & Johnson's G12C covalent inhibitor JNJ-74699157 (also known as ARS-3248) is in Phase 1 as monotherapy (NCT04006301) and Eli Lilly and Company's LY3499446 is in Phase 1/2 studies as mono- and combination therapy (NCT04165031).

From the covalent inhibitors in the clinical development, AMG 510 in complex with KRAS was the first structure that was made publicly available (Fig. 1B) (PDB ID: 6oim)¹². The compound is binding to switch-II pocket (S-IIP) and specifically exploits a smaller sub-pocket formed by H95, Y96 and Q99 with isopropyl substituent of the *p*-methylpyridine. The KRAS(G12C)–AMG 510 complex is a high-resolution structure (1.65 Å)

¹Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls University Tübingen, Tübingen, Germany. ²School of Pharmacy, University of Eastern Finland, Kuopio, Finland. email: tatu.pantsar@uef.fi

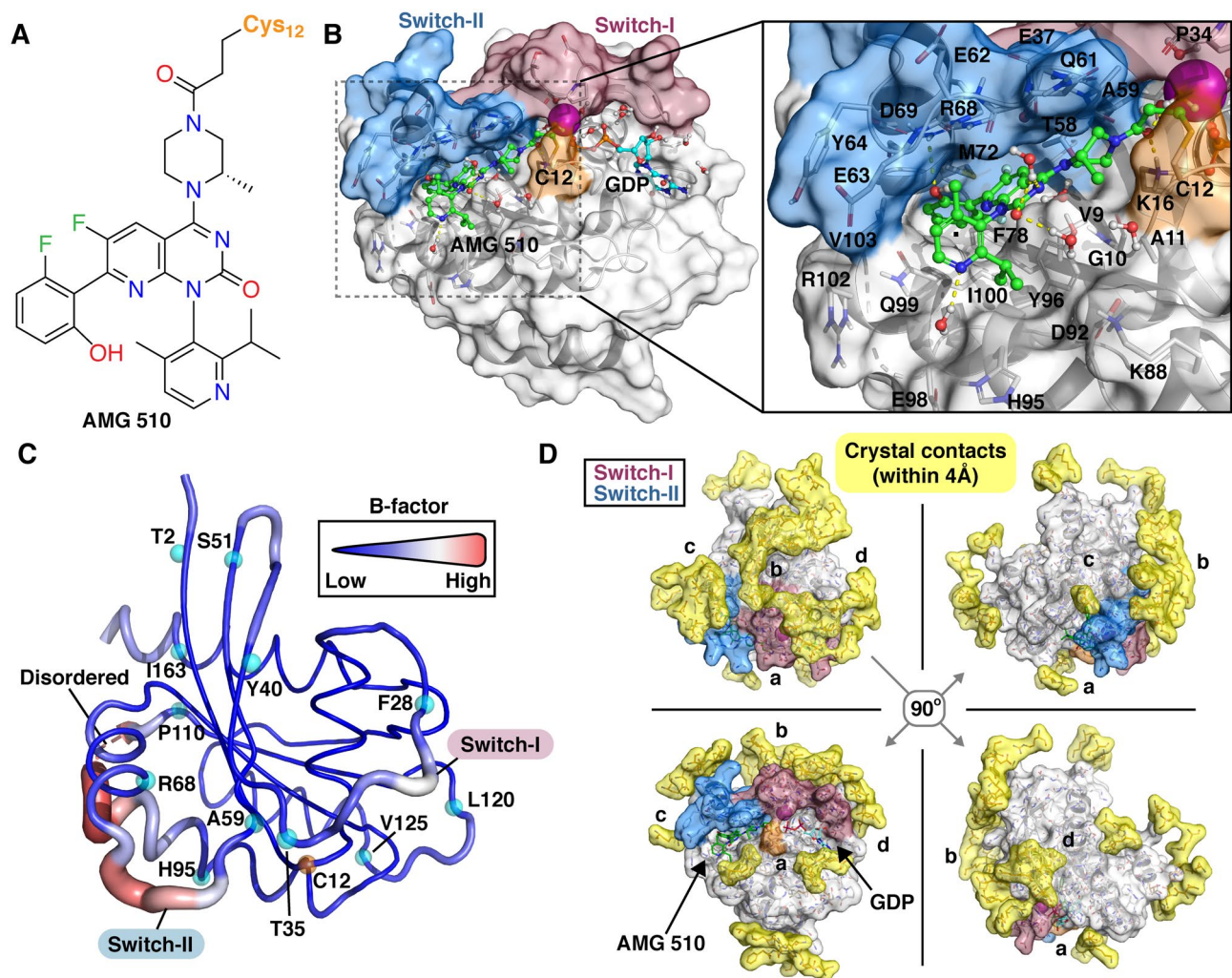


Figure 1. KRAS(G12C)–AMG 510 structure. (A) Chemical structure of AMG 510. (B) Crystal structure of KRAS(G12C)–AMG 510 complex (PDB ID: 6oim). AMG 510 (green) binds to the SII-P and exploits a cryptic pocket formed by H95/Y96/Q99. (C) *B*-factor values of KRAS(G12C)–AMG 510 structure. Thick and red ribbon indicate a high *B*-factor value, whereas thin and blue a low value. Ca-atoms of start and end residues of higher *B*-factor regions are highlighted with cyan spheres (regions: G0–T2; F28–T35; Y40–S51; A59–R68; H95–P110; L120–V125; I163–H166). A disordered region of the structure, D105–E107, is indicated with dashed red line. (D) Observed crystal contacts in KRAS(G12C)–AMG 510 structure. Residues within 4 Å of the protein are shown in sticks together with their molecular surface (yellow). In (B, D) KRAS switch regions: switch-I: residues 25–40 and switch-II: residues 58–72 are highlighted with red and blue, respectively. C12 is highlighted with orange.

where both switch regions are fully ordered. In this structure, as with majority of KRAS(G12C) structures with covalently bound ligand, all native cysteine residues are mutated (C51S; C80L; C118S)¹⁵. In fact, there is only three of KRAS(G12C) covalent inhibitor complex structures available in the Protein Data Bank without any engineered mutations (PDB IDs: 5v71¹⁶, 5v91¹⁷, 5v90¹⁷). To note, the effect of these engineered mutations to KRAS conformational dynamics is unclear to date. The KRAS(G12C)–AMG 510 structure also contains an additional glycine (G0) residue at its N-terminal.

In KRAS(G12C)–AMG 510 structure, atomic displacement parameter, also known as *B*-factor that indicates the atomic fluctuations in the crystal¹⁸, displays the highest values in switch-II and at the end of the helix- α 3, which is connected to a disordered loop region D105–E107 (Fig. 1C). Conversely, based on *B*-factor the end-region of switch-I, interswitch region and beginning of switch-II appear relatively stable. Crystal contacts, however, may have influence on switch conformations and stability, which is in fact a commonly shared characteristics among KRAS structures that display ordered switches¹⁵. Indeed, a closer inspection revealed crystal contacts on top of the stabilized switch regions (Fig. 1D). This suggest possibility that the crystal contacts may play a decisive role for the observed switch conformations in the KRAS(G12C)–AMG 510 structure.

Ligand binding may change the free energy landscape of its target protein, which defines conformations and their frequency that the protein populates¹⁹. Discrepancy in KRAS conformational dynamics for G12C targeting covalent inhibitors was revealed by hydrogen/deuterium-exchange mass spectrometry (HDX MS)¹⁶. In the study,

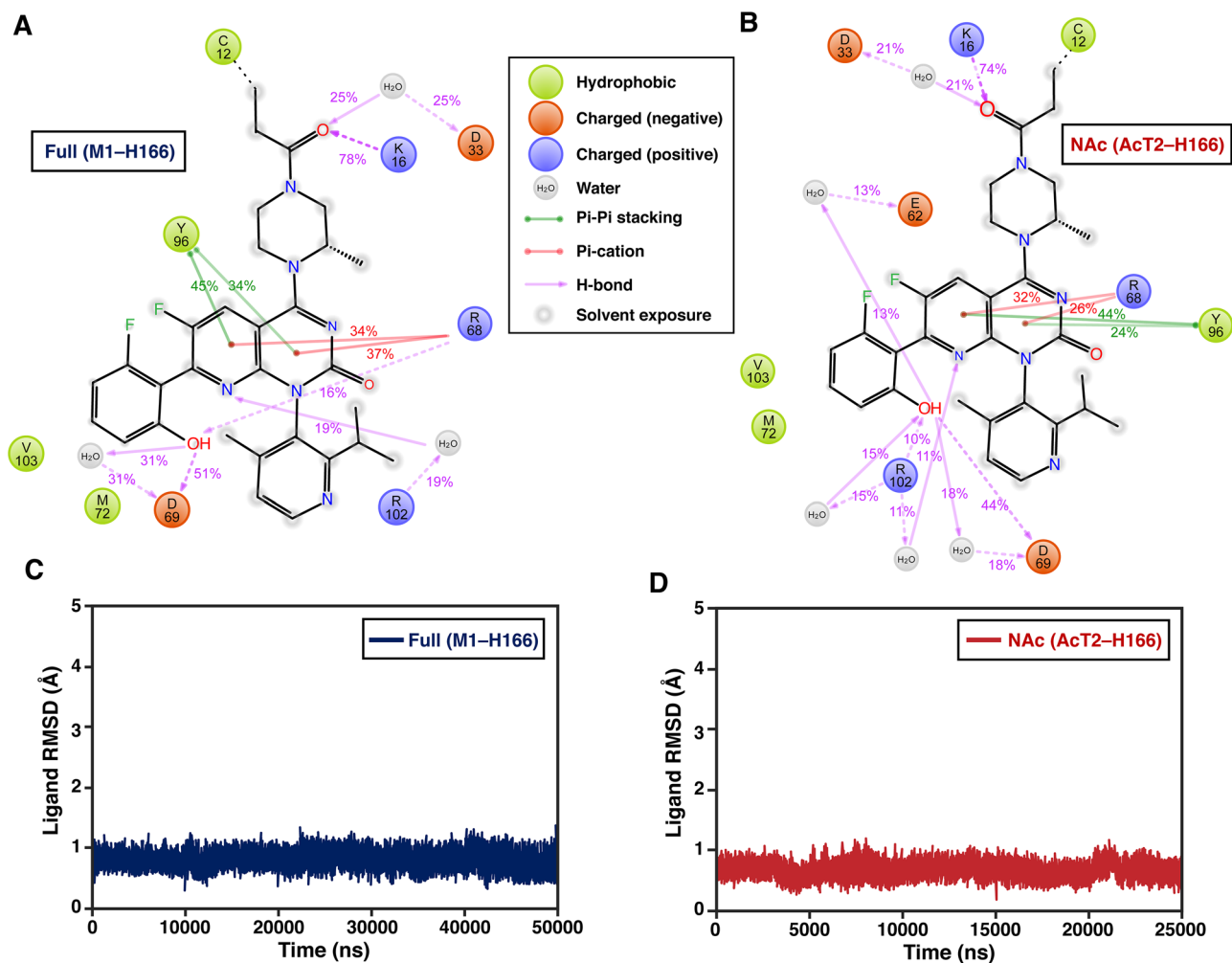


Figure 2. KRAS(G12C)-AMG 510 interactions in the MD simulations. **(A)** Observed KRAS(G12C)-AMG 510 interactions with the simulations of Full (M1-H166) systems (total of 50 μ s simulation data). **(B)** Observed KRAS(G12C)-AMG 510 interactions with the simulations of NAc (AcT2-H166) systems (total of 25 μ s simulation data). Interactions that occur more than 10% of the simulation time are displayed. **(C)** Ligand root-mean-square deviation (RMSD) in the simulations of Full (M1-H166) systems. **(D)** Ligand RMSD in the simulations of NAc (AcT2-H166) systems.

Westover et al. demonstrated that chloro-hydroxy aniline and quinazoline inhibitors exhibit unique influence on KRAS conformational dynamics. To date, however, no data has been reported of how AMG 510 influences on KRAS dynamics.

Recently, a post-translational modification (PTM) of KRAS, N acetylation of T2 with the excision of M1 (hereafter shortened as NAc) has gained more attention^{20–22}. Based on the published structures by Dharmiah et al.²⁰, acetylation of the T2 is crucial for stabilizing KRAS conformation when the initiator M1 is excised from the structure. The acetyl group displays a water mediated contact to residues located in β 2- β 3-sheets in the interswitch region. No MD simulations of KRAS with this PTM have been reported; therefore, its influence on KRAS dynamics is still somewhat unclear.

Here, we utilized long timescale MD simulations to investigate the dynamic interactions of the KRAS(G12C)-AMG 510 complex. Our simulations suggest, in contrast with what is observed in crystal structure of the KRAS(G12C)-AMG 510 complex, that the switches are not in a closed defined configuration but rather appear in an ensemble of conformations. Furthermore, we investigated for the first-time the effect of the N-terminal PTM on KRAS dynamics by MD simulations. Interestingly, our data suggests that this PTM further stabilizes KRAS in the interswitch β 2- β 3-loop compared to KRAS with M1.

Results

KRAS(G12C)-AMG 510 interactions in the dynamic environment. First, we evaluated how well KRAS(G12C)-AMG 510 interactions observed in the crystal structure reflect to the ones observed in dynamic environment (Supplementary Fig. S1; Fig. 1B). Overall, the interactions observed in the crystal structure are well maintained throughout the MD simulations of Full (M1-H166) systems (Fig. 2A). Similar interaction profile is also displayed with N-terminally modified NAc (AcT2-H166) systems (Fig. 2B). However, dissimilarities

in interactions between the crystal structure and MD simulations exist. Remarkably, methylpiperazine linker appears solvent exposed and is not shielded from water by switch-II (Fig. 2A,B), while in the crystal structure it is enclosed beneath the switch-II (Fig. 1B; Supplementary Fig. S1). In addition to the interaction with K16, the carbonyl-oxygen of AMG 510 next to the covalent attachment point displays a water bridged interaction to residue D33 in switch-I (Fig. 2A,B). This reveals a surprising possibility for AMG 510 to interact with switch-I. This interaction is displayed in two out of five replicate simulations with both systems (Supplementary Figs. S2–S3). Hydroxyl group of AMG 510, which do not present any direct polar interaction in the crystal structure, displays several possibilities for polar interactions, as water bridged interactions to E62, D69 or R102 and direct H-bonds with D69 or R102 are observed.

AMG 510 exploits so-called cryptic pocket in SII-P, which is formed by the residues H95/Y96/Q99²³. From these residues, Y96 forms π - π interactions with azaquinoxaline and these interactions are well conserved throughout the simulations (Fig. 2A,B); whereas Q99 and H95 form more unspecific water bridges (Supplementary Figs. S2–S3). Furthermore, from this heteroaromatic azaquinoxaline a cation- π interaction is occurring with R68 of switch-II.

In the energy-minimized crystal structure, carbonyl group of the azaquinoxaline displays interactions to two water molecules, one forming a putative water bridge to Q61. However, throughout the simulations no conserved interactions are observed for this carbonyl oxygen. Similarly, from the polar heteroatoms of AMG 510, aromatic nitrogens in the solvent interface are not displaying any conserved interactions during simulations, whereas pyridine nitrogen interacts with water in the crystal structure. Nevertheless, their role to interact with water is quite evident due to their location on solvent interface. These short-lived not-conserved interactions indicate that solvent is disordered in this AMG 510 interface.

Regardless of the increased solvent exposure compared to the crystal structure, AMG 510 is extremely stable throughout the simulations (Fig. 2C,D, Supplementary movies). This implies that a closed switch configuration is not decisive for the compound stability at the SII-P binding site.

AMG 510 influences on KRAS dynamics. Next, we next ascertained if AMG 510 has an influence on KRAS protein dynamics. Based on the protein backbone root-mean-square fluctuation (RMSF) values, overall dynamics of AMG 510 KRAS complex appears similar as observed previously for GDP bound KRAS in long timescale simulations² (Fig. 3). However, differences are also evident. Switch-II region fluctuations are alleviated compared to values observed for KRAS bound to GDP, indicating a stabilization of this switch by AMG 510 (Fig. 3A). Yet, this switch-II stabilization is not so evident with NAc systems (Fig. 3B). Residues forming the cryptic H95/Y96/Q99 pocket are remarkably stabilized in both systems. Moreover, AMG 510 not only stabilizes these pocket residues with direct interactions, but also the whole α 3-helix and part of the loop after the helix (residues N86–E107) (Fig. 3C). This stabilization, however, is lower in NAc systems compared to Full systems, but still exists. Also, NAc systems exhibit a trend of higher fluctuation in the end-part of switch-I (D33–S39) compared to GDP only or Full systems. In contrast, NAc systems display remarkable stability in the loop between the β 2 and β 3 sheets (β 2- β 3-loop) (Fig. 3D). This observation suggests a clear influence from this PTM to the dynamics of β 2- β 3-loop in the interswitch region.

AMG 510 do not stabilize KRAS switches in the crystal conformation. Based on the RMSF values, switch regions of KRAS appear highly flexible. Next, we further examined locations of these regions in simulations compared to the crystal structure. To this end, we monitored the distances from C12 to selected reference residues: F28; D33 and T35 in switch-I and Q61; E62 and Y64 in switch-II (Fig. 4). Compared to KRAS(G12C)-AMG 510 crystal structure, simulations indicate clearly higher distances especially in the connection interface of the switches: in end of switch-I and beginning of switch-II. This is demonstrated by higher distances of C12–T35 and C12–Q61 (Fig. 4C,D). In contrast, distances from C12 to the beginning of switch-I (F28) and to the end of switch-II (Y64) are in better agreement with the crystal structure. These results indicate that the switches clearly prefer more open conformations, particularly in the connection interface of the switches.

KRAS(G12C)-AMG 510 complex populate an ensemble of conformations. Switch regions display highly dynamical characteristics in the MD simulations that deviate from the observed KRAS(G12C)-AMG 510 crystal structure conformation. Based on the data presented above, switches prefer more open states compared to the crystal structure; however, with ambiguous configurations. To obtain more comprehensive information of their conformational distribution, we utilized Markov state modelling (MSM) approach^{24,25}. MSM is an optimal method to disclose relevant conformational states of a biomolecule from a long timescale simulation data.

MSM of Full (M1–H166) systems revealed four metastable states (S_1 – S_4) for the KRAS(G12C)-AMG 510 complex (Fig. 5). Overall, it is quite evident that all these states deviate from the crystal structure conformation (Fig. 5C). In metastable states S_1 and S_2 , switch-II is forming a helical conformation that is in orientation perpendicular to α 2-helix. This type of helical switch-II configuration is observed in many published KRAS crystal structures (Supplementary Fig. S4) (e.g. PDB IDs: 6quu²⁶; 6gj7²⁷; 6mqg²⁸; 6bof²⁹; 6qux²⁶; 6quw²⁶; 6quv²⁶). Overall these S_1 and S_2 conformations appear similar; however, with S_2 switch-I is in more open conformation at the beginning of the switch (residues D30–D33) and also α 2-helix is more stable with this state.

In the second most populated state S_3 , the end-part of switch-I (starting from T35) exists in a fully open conformation (Fig. 5C). This conformation resembles the KRAS conformation observed in SOS1–KRAS complex (Supplementary Fig. S5) (PDB ID: 6ep1)³⁰. In fact, this might be a native pre-bound configuration of switch-I for GDP-bound KRAS that is required for SOS1 binding and facilitates the nucleotide exchange. Furthermore, S_3 displays a specific helical configuration at the beginning of switch-II (D57–Q61).

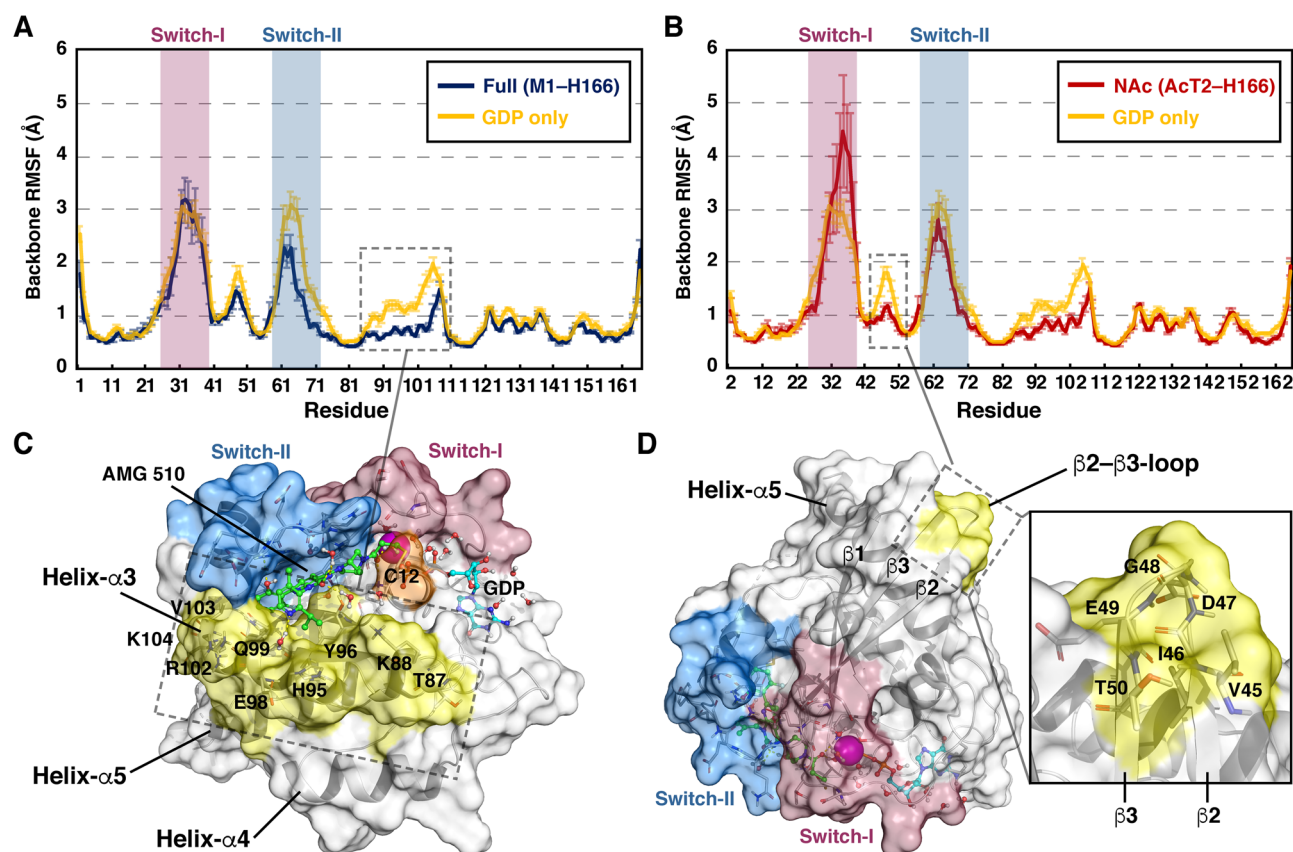


Figure 3. Root-mean-square fluctuations (RMSF) of protein backbone. (A) RMSF of full M1–H166 systems (50 μ s simulation data). (B) RMSF of NAc (AcT2–H166) systems (25 μ s simulation data). The ‘GDP only’ RMSF data in A and B is taken from previously published 70 μ s KRAS–GDP simulations². Error bars represent standard error. (C) Lower RMSF values are observed in both systems for the Helix- α 3, which location and residues are highlighted in yellow (PDB ID: 6oim). (D) Lower RMSF values are observed with NAc systems for the β 2– β 3-loop, which location and residues are highlighted in yellow (PDB ID: 6oim).

The most populated metastable state S_4 exhibits a loop-like configuration of switch-II. The switch-I conformation displays higher variability within this metastable state compared to other states, as both closed and open conformations are observed. These variable switch-I configurations are probably reflecting to the observed multiple energy minima that are included within this state (Fig. 5A).

Transitions between these metastable states occur in microsecond timescale, where mean-first-passage-time of transitions vary between 1.76 μ s (from S_3 to S_4) and 16.79 μ s (from S_4 to S_2) (Supplementary Fig. S6).

Discussion

Our simulations reveal novel atomic scale insights of a KRAS(G12C) targeting inhibitor in the clinical development. Surprisingly, AMG 510 does not fix KRAS switches to a closed conformation as would be expected based on the crystal structure. Instead, in the simulations this complex exists in an ensemble of open conformations, confirmed by MSM. In fact, KRAS(G12C)–AMG 510 complex appears to follow generally the native dynamics of GDP-bound KRAS. The observed conformations for the switches in the KRAS(G12C)–AMG 510 crystal structure are most likely induced by neighbouring proteins in the crystal environment (crystal contacts). This is indeed a common feature observed in all KRAS structures that display ordered switch conformations¹⁵. These open states of KRAS switches have been also observed in other long timescale simulations with a membrane, where a different force field was used³¹.

Regardless of the switch fluctuations, AMG 510 appears extremely stable throughout the simulations. This occurs even though the switches are not fully shielding AMG 510 from solvent. This implies that an inhibitor bound to SII-P needs to be optimized for a dynamic environment. In fact, the observed switch fluctuations that occur even when an inhibitor is bound, may provide a putative explanation for difficulties in interpretation of structure–activity-relationship (SAR) with KRAS binding ligands. Additional interesting feature is that through a water bridged interaction to D33, AMG 510 can have direct connection and influence on switch-I. This adds an additional layer of complexity in KRAS–inhibitor interactions.

Although switch regions of KRAS are highly dynamic, other parts of the protein appear more fixed. Especially the cryptic pocket that is exploited by AMG 510 is extremely stable throughout the simulations. Interestingly,

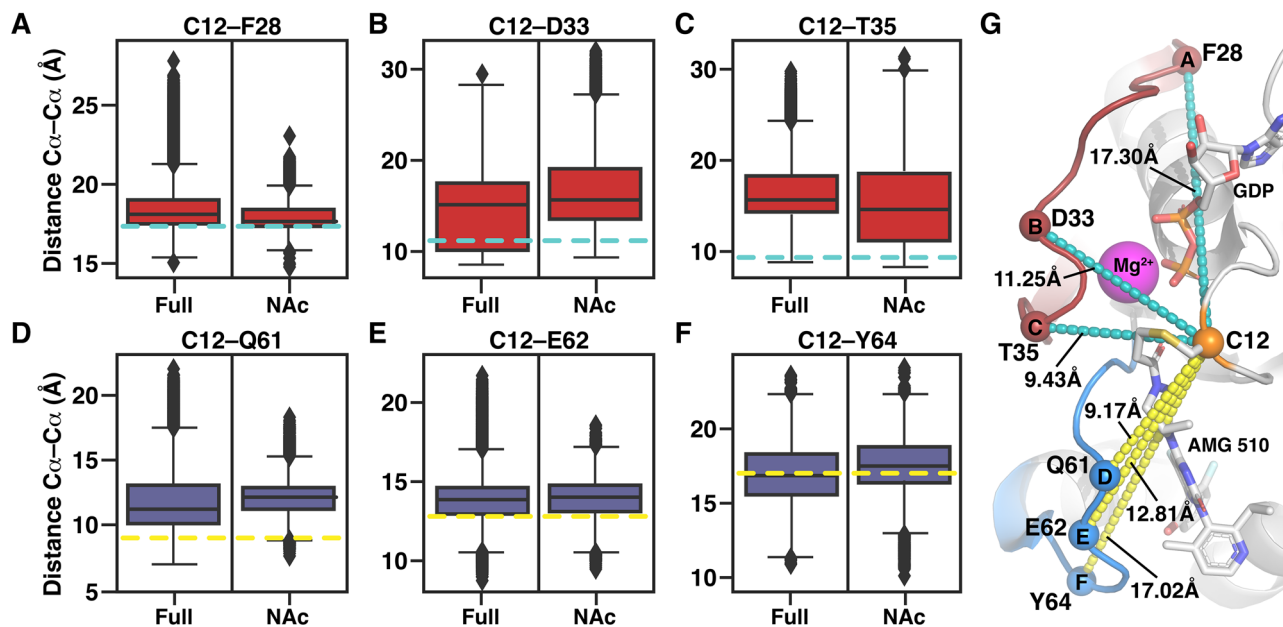


Figure 4. Distances of switch regions to selected switch-I (A–C) and switch-II (D–F) residues. The corresponding locations of the monitored Ca atoms (spheres) and their distances in the AMG 510 crystal structure (PDB ID: 6oim) are illustrated with cyan and yellow dashed lines in A–F. The black horizontal line in the box represents the median. Box displays the quartiles of the dataset (25%–75%) and whiskers the rest of the data with maximum 1.5 IQR. Outliers are indicated with black diamonds. (G) Monitored distances represented in the KRAS(G12C)–AMG 510 structure.

stabilization of this pocket by AMG 510 further stabilizes KRAS dynamics in the whole α 3-helix region. To note, this region may play an important role in KRAS dimerization and nanoclustering^{32–34}.

One metastable state of KRAS(G12C)–AMG 510, state S_3 , resembles a pre-nucleotide exchange conformational state. Although switch-II conformation in this state when an inhibitor is bound to SII-P disallows binding of SOS1, it may be important for the efficacy of an inhibitor to allow KRAS to visit these native conformations.

To the best of our knowledge, we report here for the first-time MD simulations with N-terminally processed KRAS. As shown by Dharmiah et al.²⁰, N-acetylation of T2 stabilizes KRAS dynamics upon the excision of M1. Here with AMG 510, dynamics of full N-term with iMet and N-acetylated (T2) appear also comparable, agreeing with the reported structural observations. Furthermore, the interaction profile of AMG 510 is similar with or without this PTM (Fig. 2). Therefore, N-terminal modification of KRAS is not expected to affect binding of S-IIP inhibitors, as is clearly demonstrated with the observed efficacy of AMG 510^{12,23}. Nevertheless, this PTM appears to stabilize KRAS dynamics in the loop region that connects beta-sheets β 2 and β 3 (residues 45–50). This β 2– β 3-loop may play a crucial role at KRAS dynamics at the membrane, as based on the ‘exposed’ configuration NMR data driven structure this loop exist in a direct contact to the membrane (PDB ID: 2msc)³⁵. Moreover, in a long timescale simulations study of KRAS at the membrane this orientation was frequently populated³¹. Recent NMR data driven models of KRAS–RAF on nanodiscs suggest that this loop is in contact with the CRD-domain of RAF (PDB IDs: 6pts, 6ptw)³⁶. In a study with HRAS, mutations to residues in this region D47A, E49A were shown to enhance RAS nanoclustering and its signalling activity³⁷. Also, T50I mutation with NRAS is observed in Noonan syndrome that leads to increased signalling³⁸. These results highlight the fact that it is important to take into account the possibility that N-terminal processed KRAS may behave differently at the membrane compared to full-length KRAS with M1 present, which has been used in most of the studies. Therefore, future research should clarify if this N-terminal PTM affects KRAS dynamics at the membrane.

The dynamics here generally follow what has been previously observed for KRAS. As always, in MD simulations the selected methodology could however cause bias to the observed dynamics (see details in methods). For instance, the applied force field here is not a polarizable force field³⁹, which could have influence on the results.

Overall, these results provide novel atomic-level insights to KRAS(G12C)–inhibitor complex. First, they suggest that KRAS exists in an ensemble of conformations when AMG 510 is bound. Second, they indicate that a KRAS targeting inhibitor should be optimized to a more solvent exposed dynamic environment than what a crystal structure suggests, as the switches appear dynamic even when an inhibitor is present in SII-P. Finally, potential influence of N-terminal PTM of KRAS at the membrane needs to be clarified in future. Proper understanding of KRAS and its conformational dynamics is crucial, especially when targeting other mutant forms of this oncoprotein which are lacking the cysteine for covalent inhibition.

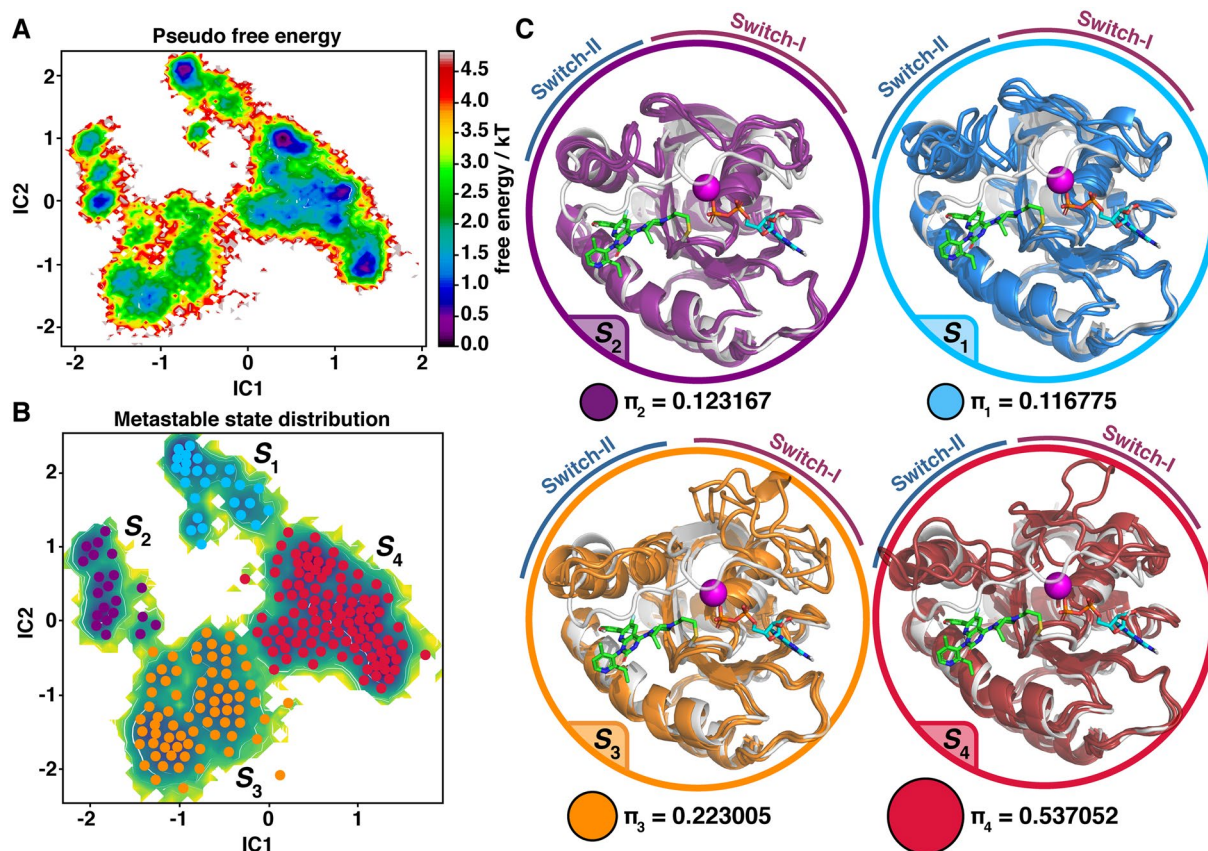


Figure 5. Metastable states of KRAS(G12C)–AMG 510 complex revealed by Markov state modelling. (A) Pseudo free energy map of distributions along time-lagged independent components (ICs) 1 and 2. (B) Separation of four metastable states (S_1 – S_4) by PCCA++ analysis. (C) Visualization of the metastable states. Each metastable state (S_i) is illustrated with three representative structures (coloured cartoons) and the crystal structure conformation of KRAS(G12C)–AMG 510 complex as a reference in grey. Equilibrium probability (π_i) for each state is indicated below the conformations together with circles with an area that is relative to state probability.

Methods

Molecular dynamics simulations. All the molecular modelling was conducted with Maestro⁴⁰ using OPLS3e force field^{41,42}. OPLS3e generated parameters were applied for the ligands, GDP and AMG 510. For the simulations, we utilized the AMG 510's lead molecule – KRAS complex (PDB ID: 6p8y⁴³). Engineered mutations of the structure were reverse mutated back to native wild-type form (C51S; C80L; C118S). The ligand was manually changed to AMG 510 and the complex was optimized and energy-minimized by Protein preparation wizard⁴⁴. Additionally, for the PTM NAC systems, M1 was deleted and the resulting terminal T2 was acetylated.

Simulations were run using Desmond MD engine⁴⁵. These minimized KRAS(G12C)–AMG 510 complexes were solvated to cubic boxes with a minimum distance of 15 Å from the protein. Water was described with TIP3P water model⁴⁶. K^+ -ions and Cl^- -ions were added to obtain 0.15 M ionic strength with a total net charge neutral. MD simulations were run in NpT ensemble ($T = 310$ K, Nosé-Hoover method; $p = 1.01325$ bar, Martyna-Tobias-Klein method) with default Desmond settings. RESPA integrator with 2 fs, 2 fs and 6 fs timesteps were used for bonded, near and far, respectively. The default value of 9 Å was used for Coulombic cutoff.

Five replicate simulation for each system using different seed numbers were run. Full (M1–H166) KRAS(G12C)–AMG 510 simulation were simulated for 10 μ s and NAC systems for 5 μ s, resulting in total of 50 μ s and 25 μ s aggregate simulations, respectively. A default Desmond relaxation protocol was applied before the production simulations.

Maestro tools were used for all of the analysis, except for MSM (see MSM details below). Visualization of the structures was done with PyMOL⁴⁷.

Markov state model generation. MSM generation was conducted with PyEMMA 2⁴⁸. Bayesian MSM was conducted following the general recommendations⁴⁹. The individual trajectories of full M1–H166 systems were used as an input for MSM generation. For featurization, we used the backbone torsions of residues from switch-I and switch-II together with selected residues in contact with AMG 510 (residues 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 95, 96, 99, 103). During the decision process of which residues were selected to be included in the final model, we monitored closely

the VAMP-2 score⁵⁰, which was 2.55 for the final model. Dimensional reduction was conducted with time-lagged independent component analysis (TICA)⁵¹. Lag time $\tau = 10$ ns and two dimensions were selected, where the implied timescales were converged, and further validation was conducted with the Chapman-Kolmogorov validation test (Supplementary Fig S7). Discretization of the data to microstates was done by k-means clustering (\sqrt{N} used for the number of clusters). Finally, a spectral clustering using the Perron-cluster cluster analysis (PCCA++)⁵² was used to assign the microstates to metastable macrostates. Transition-path theory (TPT) was applied to investigate state transitions and the flux between metastable states^{53,54}.

Data availability

The datasets generated during and analysed during the current study are freely available in the Zenodo repository, <https://doi.org/10.5281/zenodo.3711537>. The dataset include original raw-trajectories, movies of the MD simulations and PDB-coordinates for the metastable states.

Received: 18 March 2020; Accepted: 3 July 2020

Published online: 20 July 2020

References

- Mo, S. P., Coulson, J. M. & Prior, I. A. RAS variant signalling. *Biochem. Soc. Trans.* **46**, 1325–1332 (2018).
- Pantsar, T. *et al.* Assessment of mutation probabilities of KRAS G12 missense mutants and their long-timescale dynamics by atomistic molecular simulations and Markov state modeling. *PLoS Comput. Biol.* **14**, e1006458 (2018).
- Prior, I. A., Lewis, P. D. & Mattos, C. A comprehensive survey of Ras mutations in cancer. *Cancer Res.* **72**, 2457–2467 (2012).
- Dogan, S. *et al.* Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin. Cancer Res.* **18**, 6169–6177 (2012).
- Román, M. *et al.* KRAS oncogene in non-small cell lung cancer: Clinical perspectives on the treatment of an old target. *Mol. Cancer* **17**, 33 (2018).
- Dang, C. V., Reddy, E. P., Shokat, K. M. & Soucek, L. Drugging the ‘undruggable’ cancer targets. *Nat. Rev. Cancer* **17**, 502–508 (2017).
- Gehring, M. & Laufer, S. A. Emerging and re-emerging warheads for targeted covalent inhibitors: Applications in medicinal chemistry and chemical biology. *J. Med. Chem.* **62**, 5673–5724 (2019).
- Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
- Patricelli, M. P. *et al.* Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer Discov* **6**, 316–329 (2016).
- Janes, M. R. *et al.* Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor. *Cell* **172**, 578–589.e17 (2018).
- Goody, R. S., Müller, M. P. & Rauh, D. Mutant-specific targeting of Ras G12C activity by covalently reacting small molecules. *Cell Chem. Biol.* **26**, 1338–1348 (2019).
- Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223. <https://doi.org/10.1038/s41586-019-1694-1> (2019).
- Hallin, J. *et al.* The KRASG12C inhibitor, MRTX849, provides insight toward therapeutic susceptibility of KRAS mutant cancers in mouse models and patients. *Cancer Discov.* <https://doi.org/10.1158/2159-8290.CD-19-1167> (2019).
- AMG 510 activity in subjects with advanced solid tumors with KRAS p.G12C mutation (CodeBreak 101)—Full Text View—ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT04185883>.
- Pantsar, T. The current understanding of KRAS protein structure and dynamics. *Comput. Struct. Biotechnol. J.* **18**, 189–198 (2020).
- Lu, J. *et al.* KRAS G12C drug development: Discrimination between switch II pocket configurations using hydrogen/deuterium-exchange mass spectrometry. *Structure* **25**, 1442–1448.e3 (2017).
- Zeng, M. *et al.* Potent and selective covalent quinazoline inhibitors of KRAS G12C. *Cell Chem. Biol.* **24**, 1005–1016.e3 (2017).
- Trueblood, K. N. *et al.* Atomic displacement parameter nomenclature. Report of a subcommittee on atomic displacement parameter nomenclature. *Acta Cryst. A* **52**, 770–781 (1996).
- Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
- Dharmaiah, S. *et al.* Structures of N-terminally processed KRAS provide insight into the role of N-acetylation. *Sci. Rep.* **9**, 10512 (2019).
- Ntai, I. *et al.* Precise characterization of KRAS4b proteoforms in human colorectal cells and tumors reveals mutation/modification cross-talk. *PNAS* **115**, 4140–4145 (2018).
- Buser, C. A. *et al.* High-performance liquid chromatography/mass spectrometry characterization of Ki4B-Ras in PSN-1 cells treated with the prenyltransferase inhibitor L-778,123. *Anal. Biochem.* **290**, 126–137 (2001).
- Lanman, B. A. *et al.* Discovery of a covalent inhibitor of KRASG12C (AMG 510) for the treatment of solid tumors. *J. Med. Chem.* **63**, 52–65 (2020).
- Chodera, J. D. & Noé, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **25**, 135–144 (2014).
- Husic, B. E. & Pande, V. S. Markov state models: From an art to a science. *J. Am. Chem. Soc.* **140**, 2386–2396 (2018).
- Bergner, A. *et al.* KRAS binders hidden in nature. *Chemistry* **25**, 12037–12041 (2019).
- Kessler, D. *et al.* Drugging an undruggable pocket on KRAS. *Proc. Natl. Acad. Sci. USA* **116**, 15823–15829 (2019).
- Bera, A. K. *et al.* Structural basis of the atypical activation mechanism of KRASV14I. *J. Biol. Chem.* **294**, 13964–13972 (2019).
- Poulin, E. J. *et al.* Tissue-specific oncogenic activity of KRASA146T. *Cancer Discov.* **9**, 738–755 (2019).
- Hillig, R. C. *et al.* Discovery of potent SOS1 inhibitors that block RAS activation via disruption of the RAS-SOS1 interaction. *Proc. Natl. Acad. Sci. USA* **116**, 2551–2560 (2019).
- Neale, C. & Garcia, A. E. The plasma membrane as a competitive inhibitor and positive allosteric modulator of KRAS4B signaling. *Biophys. J.* <https://doi.org/10.1016/j.bpj.2019.12.039> (2020).
- Sarkar-Banerjee, S. *et al.* Spatiotemporal analysis of K-Ras plasma membrane interactions reveals multiple high order homo-oligomeric complexes. *J. Am. Chem. Soc.* **139**, 13466–13475 (2017).
- Muratcioglu, S. *et al.* GTP-dependent K-Ras dimerization. *Structure* **23**, 1325–1335 (2015).
- Muratcioglu, S. *et al.* Oncogenic K-Ras4B dimerization enhances downstream mitogen-activated protein kinase signaling. *J. Mol. Biol.* **432**, 1199–1215 (2020).
- Mazhab-Jafari, M. T. *et al.* Oncogenic and RASopathy-associated K-RAS mutations relieve membrane-dependent occlusion of the effector-binding site. *Proc. Natl. Acad. Sci. USA* **112**, 6625 (2015).

36. Fang Z, *et al.* Multivalent assembly of KRAS with the RAS-binding and cysteine-rich domains of CRAF on the membrane. *Proc Natl Acad Sci U S A* **117**(22), 12101–12108 (2020).
37. Šolman, M. *et al.* Specific cancer-associated mutations in the switch III region of Ras increase tumorigenicity by nanocluster augmentation. *eLife* **4**, e08905 (2015).
38. Cirstea, I. C. *et al.* A restricted spectrum of NRAS mutations causes Noonan syndrome. *Nat Genet* **42**, 27–29 (2010).
39. Jing, Z. *et al.* Polarizable force fields for biomolecular simulations: Recent advances and applications. *Annu. Rev. Biophys.* **48**, 371–394 (2019).
40. *Schrödinger Release 2019–3: Maestro*, Schrödinger, LLC, New York, NY (2019).
41. Roos, K. *et al.* OPLS3e: Extending force field coverage for drug-like small molecules. *J. Chem. Theory Comput.* **15**, 1863–1874 (2019).
42. Harder, E. *et al.* OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **12**, 281–296 (2016).
43. Shin, Y. *et al.* Discovery of N-(1-acryloylazetid-3-yl)-2-(1H-indol-1-yl)acetamides as covalent inhibitors of KRASG12C. *ACS Med. Chem. Lett.* **10**, 1302–1308 (2019).
44. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).
45. Bowers, K. J. *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* (ACM, 2006). <https://doi.org/10.1145/1188455.1188544>.
46. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
47. Schrödinger. *The PyMOL Molecular Graphics System, Version 2.0* (2019).
48. Scherer, M. K. *et al.* PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *J. Chem. Theory Comput.* **11**, 5525–5542 (2015).
49. Wehmeyer, C. *et al.* Introduction to Markov state modeling with the PyEMMA software [Article v1.0]. *Living Journal of Computational Molecular Science* **1**, 5965– (2018).
50. Wu, H. & Noé, F. Variational approach for learning Markov processes from time series data. [arXiv:1707.04659](https://arxiv.org/abs/1707.04659) [math, stat] (2019).
51. Pérez-Hernández, G., Paul, F., Giorgino, T., De Fabritiis, G. & Noé, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **139**, 015102 (2013).
52. Röblitz, S. & Weber, M. Fuzzy spectral clustering by PCCA+: Application to Markov state models and data classification. *Adv. Data Anal. Classif.* **7**, 147–179 (2013).
53. Metzner, P., Schütte, C. & Vanden-Eijnden, E. Transition path theory for Markov jump processes. *Multiscale Model. Simul.* **7**, 1192–1219 (2009).
54. Noé, F., Schütte, C., Vanden-Eijnden, E., Reich, L. & Weikl, T. R. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. USA* **106**, 19011–19016 (2009).

Acknowledgements

CSC-IT Center for Science Ltd. (Espoo, Finland) is acknowledged for computational resources. The author acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 839230 and Orion Research Foundation sr. Also, support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen is acknowledged.

Author contributions

T.P. wrote the manuscript and conducted all the research and analysis.

Competing interests

The author declares no competing interest.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-68950-y>.

Correspondence and requests for materials should be addressed to T.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020