Current Research in Structural Biology 7 (2024) 100115

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Research Article

Studying early structural changes in SOS1 mediated KRAS activation mechanism

Kirti Bhadhadhara, Vinod Jani, Shruti Koulgi, Uddhavesh Sonavane, Rajendra Joshi *

High Performance Computing-Medical & Bioinformatics Applications Group, Centre for Development of Advanced Computing (C-DAC), Innovation Park, Panchawati, Pashan, Pune, 411008, India

ARTICLE INFO

Handling Editor: Prof G Oliva

Keywords:
Molecular dynamics simulation
KRAS
SOS1
MMPBSA
Interaction
PCA

ABSTRACT

KRAS activation is known to be modulated by a guanine nucleotide exchange factor (GEF), namely, Son of Sevenless1 (SOS1). SOS1 facilitates the exchange of GDP to GTP thereby leading to activation of KRAS. The binding of GDP/GTP to KRAS at the REM/allosteric site of SOS1 regulates the activation of KRAS at CDC25/catalytic site by facilitating its exchange. Different aspects of the allosteric activation of KRAS through SOS1 are still being explored. To understand the SOS1 mediated activation of KRAS, molecular dynamics simulations for a total of nine SOS1 complexes (KRAS-SOS1-KRAS) were performed. These nine systems comprised different combinations of KRAS-bound nucleotides (GTP/GDP) at REM and CDC25 sites of SOS1. Various conformational and thermodynamic parameters were analyzed for these simulation systems. MMPBSA free energy analysis revealed that binding at CDC25 site of SOS1 was significantly low for GDP-bound KRAS as compared to that of GTP-bound KRAS. It was observed that presence of either GDP/GTP bound KRAS at the REM site of SOS1 affected the activation related changes in the KRAS present at CDC25 site. The conformational changes at the catalytic site of SOS1 resulting from GDP/GTP-bound KRAS at the allosteric changes may hint at KRAS activation through different pathways (slow/fast/rare). The allosteric effect on activation of KRAS at CDC25 site may be due to conformations adopted by switch-I, switch-II, beta2 regions of KRAS at REM site. The effect of structural rearrangements occurring at allosteric KRAS may have led to increased interactions between SOS1 and KRAS at both the sites. The SOS1 residues involved in these important interactions with KRAS at the REM site were R694, S732 and K735. Whereas the ones interacting with KRAS at CDC25 site were S807, W809 and K814. This may suggest the crucial role of these residues in guiding the allosteric activation of KRAS at CDC25 site. The conformational shifts observed in the switch-I, switch-II and alpha3 regions of KRAS at CDC25 site may be attributed to be a part of allosteric activation. The binding affinities, interacting residues and conformational dynamics may provide an insight into development of inhibitors targeting the SOS1 mediated KRAS activation.

1. Introduction

RAS protein belongs to the GTPase family of proteins and plays an important role in various cellular activities like cell proliferation, survival and differentiation (Muñoz-Maldonado et al., 2019; Lukman et al., 2010; Pantsar, 2020; Lu et al., 2016; Murugan et al., 2019). RAS is known to have three isoforms namely HRAS, NRAS and KRAS. The differences lie in the hyper-variable region and the allosteric lobe of the G-domain. Each isoform exhibits a different mutation frequency at each codon which may lead to distinct biological phenomenon (Muñoz-Maldonado et al., 2019). Among the three isoforms, KRAS is found to be more oncogenic, having a share of around 85% mutations leading to human cancers and rasopathies (Lukman et al., 2010; Pantsar, 2020). Several cellular pathways are affected owing to mutations in the KRAS protein leading to cancerous conditions (Liu et al., 2019; Gurung and Bhattacharjee, 2015; Gimple and Wang, 2019). KRAS is a 188 amino acids long protein, consisting of five α-helices and six β-strands that are responsible for forming the protein core. The details of the different regions of protein have been given in the Supplementary Table 1 (Berman et al., 2003), (Walker et al., 1982). The KRAS structure consists of a

Abbreviations: RalGDS, Ral guanine nucleotide dissociation stimulator; RAF-kinase, Rapidly Accelerated Fibrosarcoma; PI3K, phosphoinositide 3-kinase; TIAM1, T-lymphoma invasion and metastasis; RTK, receptor tyrosine kinase; TIP3P, transferable intermolecular potential with 3 points.

* Corresponding author.
E-mail address: rajendra@cdac.in (R. Joshi).

https://doi.org/10.1016/j.crstbi.2023.100115
Received 17 August 2023; Received in revised form 28 November 2023; Accepted 29 November 2023
Available online 9 December 2023

central G-domain and C-terminal region named as hyper variable region (HVR) (Pantsar, 2020). The G-domain contains the P-loop and switch regions, which form the important functional components of KRAS known to be involved in the activation process (Saraste et al., 1990). The P-loop region positions within the residue range 10–17 (Pantsar, 2020). The P-loop region is also known as Walker A motif (Lukman et al., 2010), (Saraste et al., 1990). The switch region comprises two important regions namely, switch-I (residues 25–40) and switch-II (residues 58–74) (Fig. 1B) (Parker et al., 2022). The switch regions form the binding interface for the effector proteins which act as controllers for interacting with guanosine exchange factors (GEFs)/GTPase activating protein (GAPs). These proteins are involved in the activation/deactivation of the KRAS protein (Gasper and Wittinghofer, 2020). These switch regions are in close proximity with each other and are connected via a β-sheet, namely β2-β3, which is known to play a role in the stabilization of the protein structure (Haza et al., 2021). The HVR (residue K167-M188) helps in anchoring RAS to the membrane. The P-loop and switch regions consist of residues that are also known to be mutation hotspots in most of the human cancers (Parker et al., 2022), (Chen et al., 2013). The exchange of GDP to GTP, which leads to activation of KRAS is guided by a class of proteins named guanosine exchange factors (GEFs). One such GEF is the protein Son of Sevenless 1 (SOS1). SOS exists in the form of two homologs namely, SOS1 and SOS2. SOS1 is a GEF that activates KRAS, so called as the pacemaker of KRAS (Fig. 1C). Reverting back to the inactive GDP-bound state is achieved either by the intrinsic GTPase activity of KRAS or by employing GTPase activating proteins (GAPs) (Fig. 1A). KRAS has an intrinsic GTPase activity, which is enhanced on interaction of KRAS with GAP (Moghadamchargari et al., 2021). The activation of KRAS is guided by its interaction with SOS1 (Jeng et al., 2012), (Kessler et al., 2021). Inactivation of GTP-bound active KRAS occurs upon interaction with GAP (Fig. 1A) (Liu et al., 2019). In its active state the KRAS tends to activate several downstream proteins viz. RalGDS, RAF-kinase, PI3K and others (Gimple and Wang, 2019), (Wang et al., 2021). Activation of these proteins leads to activation of RAF-ERK, PI3K-AKT, TIAM1-RAC signaling cycle leading to oncogenic transcription, cell survival, cell growth-metabolism and cell motility-migration respectively (Lukman et al., 2010), (Walker et al., 1982), (Saraste et al., 1990), (Gasper and Wittinghofer, 2020). The RTK-RAS signaling pathway is one of the most altered pathways in cancer (Gimple and Wang, 2019). SOS1 is a multidomain protein of 1333 amino acids, comprising of histone folds (HF) (residues 1–198), Dbl homology domain (DH) (residues 200–390), Pleckstrin homology domain (PH) domain (residues 444–548), Ras exchanger motif (REM) (residues 567–741), CDC25 domain (residues 780–1019), SH3 binding motif (1020–1333) and proline rich (PR) domain (Uras et al., 2020; Bandaru et al., 2019; López et al., 2021). The N-terminal and C-terminal regions of SOS1 are bonded with a helical linker. The PR domain of SOS1 helps it in bringing to the cell membrane by binding with Grb2 protein of the SH3 domain. The PH and HF domain strengthen the membrane anchoring. SOS1 has two binding sites for the KRAS namely, allosteric binding (REM domain) and catalytic binding site (CDC25 domain) (He et al., 2022). The allosteric site functions as a positive feedback loop by activating SOS1 catalytic site upon RAS binding (He et al., 2022). At the allosteric site, GTP-bound KRAS is known to bind stronger than GDP-bound KRAS (Sondermann et al., 2004), (Hoang et al., 2021). Oncogenic mutations at both KRAS and SOS1 are responsible for cancer (Gasper and Wittinghofer, 2020), (Uras et al., 2020). It is known from previous studies that binding of the GTP-bound KRAS complex at the allosteric site of SOS1 promotes the activation of KRAS by facilitating exchange of GDP to GTP (Liu et al., 2019), (He et al., 2022). The molecular dynamics (MD) simulations carried out by Liao et al. suggested that the SOS1 may mediate the activation of the KRAS by three different cycles (Kessler et al., 2021), (Liao et al., 2018). These three cycles defined by them are rare activation cycle, limited activation cycle and fast activation cycle and are defined based on binding of nucleotides (GDP/GTP) to KRAS at allosteric and catalytic sites (Liao et al., 2018).

KRAS is either present in the guanosine diphosphate (GDP) or guanosine triphosphate (GTP) bound form. KRAS is known to be in an active state when bound to GTP and in the inactive state when bound with GDP (Liu et al., 2019). Owing to the fast exchange of GTP the crystal structures have been solved in presence of its analogues namely GCP, GNP and GSP (Pantsar, 2020). The interaction between SOS1 and KRAS plays an important role in transmitting allosteric signals and is also seen from various experimental studies on KRAS, SOS1, KRAS-SOS1 where researchers have tried to identify interaction networks between these proteins (Moghadamchargari et al., 2021), (Hoang et al., 2021),

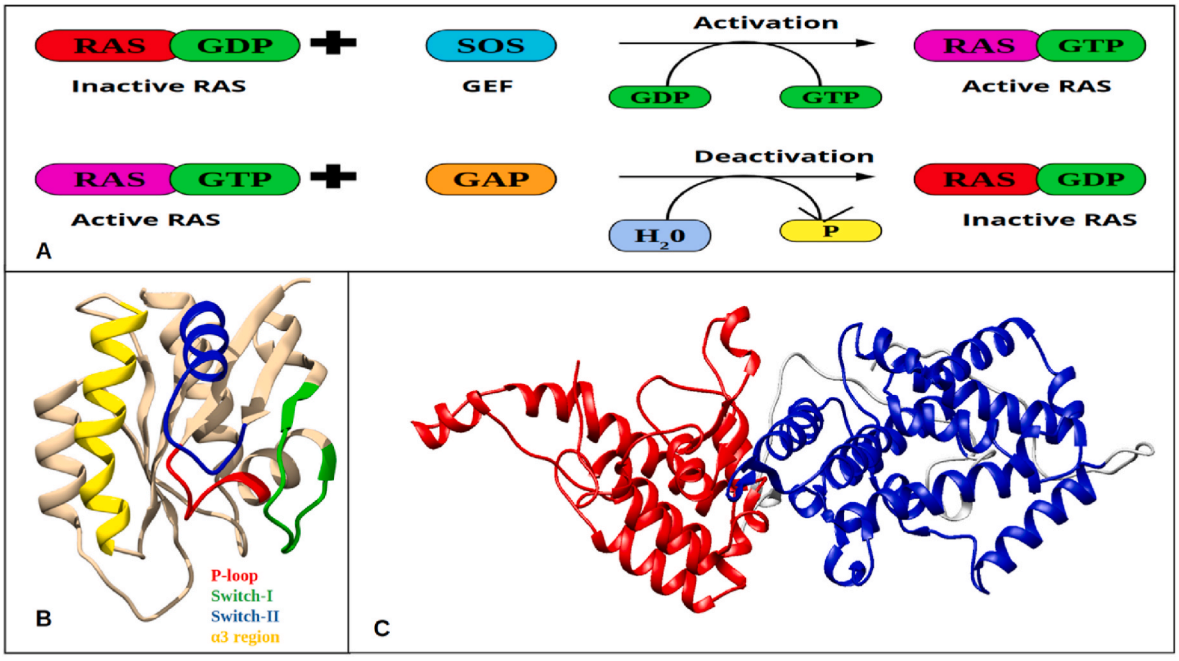

Fig. 1. (A): KRAS activation and deactivation cycle; **(B)** Structure of KRAS; **(C)** Structure of SOS1 depicting the REM site (red) and CDC25 site (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Sharma et al., 2020), (Vo et al., 2016). Besides experimental studies, various *in silico* studies have also been performed to understand the SOS1 mediated KRAS activation and network interactions responsible in the activation process (Vo et al., 2016; https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga; Vatansever et al., 2020; Iversen et al., 2014; Chen et al., 2021). Researchers had studied both KRAS and SOS1 individually and their intrinsic interaction network, but details of SOS1 mediated KRAS activation mechanism are still not clear (Maurer et al., 2012), (Sun et al., 2012). Hence, in the current study an attempt is made to understand the structural changes involved in the SOS1 mediated KRAS activation using molecular dynamics (MD) simulations. Total of nine systems have been studied, of which four systems were ternary complexes consisting of two KRAS and single SOS1, four were dimers systems consisting of single KRAS and single SOS1 and one was monomer system consisting of only SOS1 monomer (Fig. 2B). Targeting the SOS1 and KRAS interface in order to block the KRAS activation is considered one of the useful strategies in KRAS mediated cancer pathways (Ketcham et al., 2022), (Gebregiworgis et al., 2021). Therefore, understanding the KRAS-SOS1 interface in terms of the interactions formed between the two proteins may aid the researchers in the development of drugs to inhibit the SOS1 mediated KRAS activation (Ketcham et al., 2022), (Gebregiworgis et al., 2021).

2. Materials and methods

2.1. System preparation

The SOS1 systems were built in the form of dimer and ternary complexes with KRAS in either the GTP or the GDP bound form. The coordinates for SOS1 were obtained from the PDB ID: 1XD2, a ternary complex structure of SOS1 bound to nucleotide-free HRAS at the allosteric site (REM site) and GDP-bound HRAS at the catalytic site (CDC25 site) (Berman et al., 2003), (Sondermann et al., 2004). There were few missing residues in the SOS1 of 1XD2, hence they were modeled using the Swiss-Model server (Waterhouse et al., 2018). The templates and the corresponding GMQE score obtained through the Swiss-Model server have been given in Supplementary Table 2. GMQE score lies between 0 and 1 and the templates scoring close to 1 are considered as suitable for model building (Waterhouse et al., 2018). The top two templates based on the GMQE score were further selected for model building

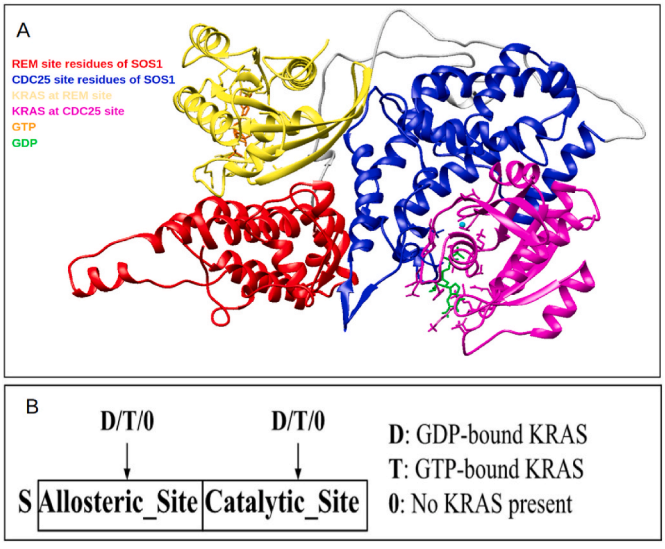

Fig. 2. (A): Structure showing the ternary complex of $KRAS^{REM}$-SOS1-$KRAS^{CDC25}$;
(B) The naming convention used for different complexes of SOS1 bound to KRAS.

(Supplementary Table 2). This modeled SOS1 was further used for building the dimer and the ternary complexes with KRAS. The conformation for GTP and GDP bound KRAS were obtained from the PDB IDs: 4DSO and 4EPT respectively (Maurer et al., 2012), (Sun et al., 2012). 4DSO and 4EPT consisted of G12D and C118S mutations, these were reverted back to the wild type residues for the KRAS. In case of dimer complexes, the KRAS at either the REM/CDC25 site of SOS1 or both were positioned by superimposing on the HRAS at the REM/CDC25 site of the PDB ID: 1XD2. A total of nine SOS1 systems were constructed namely S00, SD0, ST0, S0D, S0T, SDT, STT, STD and SDD (Fig. 2B). Here the naming convention followed for the systems depicts SOS1 protein with KRAS bound to GDP or GTP, for example system STD represents SOS1 protein bound to GTP bound KRAS at the REM site of SOS1 and GDP bound KRAS at CDC25 site of SOS1 (Fig. 2A). Similarly, the naming convention given in Fig. 2B has been followed to represent all the systems.

2.2. MD simulations

MD simulations were performed using Gromacs (v2020.3) simulation package. The parameters for the GDP and GTP were generated using the Prodrg server (Schlagenhauf et al., 2010). The AMBER FF99SB force field was used for parameterization of the proteins (Páll et al., 2020). The TIP3P water model was used for solvation. The systems were neutralized by adding counter ions Na^+ and Cl^-. The energy minimization was done for 50,000 steps using the steepest descent and conjugate gradient minimization algorithm. NVT equilibration was done for 100 ps and at 300 K followed by NPT equilibration for 100 ps. Production MD runs of 500 ns of each system were done using V-rescale for temperature coupling and Parrinello-Rahman was for pressure coupling. The time step of 2 fs was used and the short-range van der Waals cut-off of 12 Å was used. Molecular dynamics simulations were performed for each of the above mentioned nine KRAS-SOS1 systems for 500 ns.

2.3. Analysis

Analysis utilities of Gromacs and CPPTRAJ modules of Ambertools20 were used for the analysis (Páll et al., 2020; Case et al., 2021). Different structural analysis viz. root mean square deviation (RMSD), root mean square fluctuation (RMSF) and important distances were calculated (Zhao et al., 2021), (Sayyed-Ahmad et al., 2017). The visualization and rendering was performed using VMD and UCSF Chimera (Fernandes et al., 2019), (Pettersen et al., 2004). Detailed analysis of various interactions was performed using the GetContacts tool (Venkatakrishnan et al., 2019, https://getcontacts.github.io/). Xmgrace tool was used for plotting the graphs (https://plasma-gate.weizmann.ac.il/pub/grace/). Principal component analysis (PCA) was performed using the CPPTRAJ module of Ambertools20 (Case et al., 2021). The molecular mechanics poisson-boltzmann surface area (MMPBSA) free energy calculations for all nine systems were performed using the gmx_MMPBSA tool (Valdés-Tresanco et al., 2021; Miller et al., 2012; Chong et al., 2009).

3. Results

The SOS1 system is known to regulate the activation of KRAS bound to the CDC25 site (catalytic site) allosterically, owing to the presence of KRAS at its REM site (allosteric site) (He et al., 2022). The presence of KRAS in complexes with either GDP or GTP bound at the allosteric sites is known to affect the activation rate of the KRAS at the catalytic site (Liao et al., 2018). Hence, in order to understand how the presence of the KRAS bound to different nucleotides at an allosteric site guides the activation process, a total of nine systems comprising of KRAS-SOS1-KRAS ternary as well as KRAS-SOS1 dimer in the presence of GDP and/or GTP have been studied. The KRAS at REM site and CDC25 site have been referred as $KRAS^{REM}$ and $KRAS^{CDC25}$, respectively throughout the article.

3.1. Conformational variability of SOS1 and KRAS

The conformational variability of SOS1 and KRAS was determined by calculating the root mean square deviation (RMSD) for each of the systems. The start structure of the simulation was considered as the reference. The RMSD was calculated for each of the proteins in the complex namely SOS1, $KRAS^{REM}$ and $KRAS^{CDC25}$ individually for all the nine systems.

The RMSD distribution for each of the systems has been represented as the boxplots. Fig. 3A shows RMSD distribution for the SOS1 protein in each of the systems. The S00 system which consisted of only the SOS1 was observed to fluctuate the maximum with median RMSD value ranging around 5.5 Å. The dimer systems with only $KRAS^{CDC25}$ were also observed to have median RMSD values on the higher side around 4.5 and 5.3 Å for S0T and S0D respectively. The systems with only $KRAS^{REM}$ were observed to have median RMSD values around 4.7 and 3.5 Å for SD0 and ST0 respectively. The SOS1 in the systems with KRAS occupied in both REM and CDC25 sites was observed to fluctuate within 3–4.5 Å. Amongst them, the SOS1 in the SDD system was observed to fluctuate the most.

Fig. 3B shows the RMSD distribution for the $KRAS^{REM}$, switch-I and switch-II domains for the systems containing KRAS at the REM site of SOS1. In the case of the complete $KRAS^{REM}$, SD0 and ST0 were observed to fluctuate with median RMSD value ranging around 3.1 and 2.0 Å respectively. The median RMSD value was observed to be maximum in the former one. The systems having both the $KRAS^{REM}$ and $KRAS^{CDC25}$ sites occupied showed $KRAS^{REM}$ to fluctuate within the range 2.0–2.5 Å. Amongst them, the $KRAS^{REM}$ in the ST0 system was observed to fluctuate the least. The switch-I and switch-II of $KRAS^{REM}$ were observed to fluctuate the most in the SDD system with median RMSD value of 5.2 Å and 4.8 Å respectively. However, their RMSD values in the rest of the systems lies within 2.5–4.8 Å for both the switch-I and II regions.

Fig. 3C shows RMSD distribution for the complete $KRAS^{CDC25}$, switch-I and switch-II domains for the systems containing KRAS at the CDC25 site of SOS1. $KRAS^{CDC25}$ fluctuated the most and the least in the S0T and S0D with median RMSD values ranging around 5.6 Å and 2.5 Å respectively. RMSD values within 2.5–4 Å were observed for $KRAS^{CDC25}$ in the KRAS-SOS1-KRAS ternary complexes. The switch-I of $KRAS^{CDC25}$ fluctuated the most in the systems where GDP was bound to the $KRAS^{CDC25}$ with median RMSD value of 5.7 Å. The systems with GTP bound to the $KRAS^{CDC25}$ showed RMSD ranging within 2.5–5 Å. The switch-II region of $KRAS^{CDC25}$ fluctuated the most with a median RMSD value of 5 Å in the SDD system. The S0T and SDT were observed to fluctuate the least with an RMSD value of 2.5 Å.

The flexibility of switch-II of $KRAS^{CDC25}$ was observed to be slightly decreased as compared to that of switch-I in $KRAS^{CDC25}$ site for all systems. The $KRAS^{REM}$ appeared to be more rigid than the $KRAS^{CDC25}$. It was observed that the presence of the $KRAS^{REM}$ tends to show more fluctuation when bound to GDP, however the $KRAS^{CDC25}$ showed more fluctuation when bound to GTP.

3.2. Modulation of KRAS-SOS1 association

Earlier studies suggest that presence of specific nucleotides bound to the $KRAS^{REM}$ may allosterically regulate the $KRAS^{CDC25}$ of SOS1 (He et al., 2022), (Liao et al., 2018), (Hillig et al., 2019), (Margarit et al., 2003). Hence to study the changes in the proximity of KRAS with respect to SOS1, the distance between the center of mass of SOS1 and the KRAS at both the sites was calculated (Fig. 3D).

The average distance between $KRAS^{REM}$ and SOS1 lies in the range of 27.8–30.2 Å. The SD0 system was observed to have a maximum distance of around 30.2 Å. This distance was observed to be less in the presence of GTP as compared to the GDP bound $KRAS^{REM}$. The distance between $KRAS^{REM}$ and SOS1, in case of the ternary complex systems laid within the range of 28.2–29.8 Å. The average distance between $KRAS^{CDC25}$ and SOS1 lies within the range of 25–27.5 Å. The S0D system was observed to have a distance of around 26.8 Å. When the GDP in KRAS at the CDC25 site was replaced with GTP, this distance increased to around

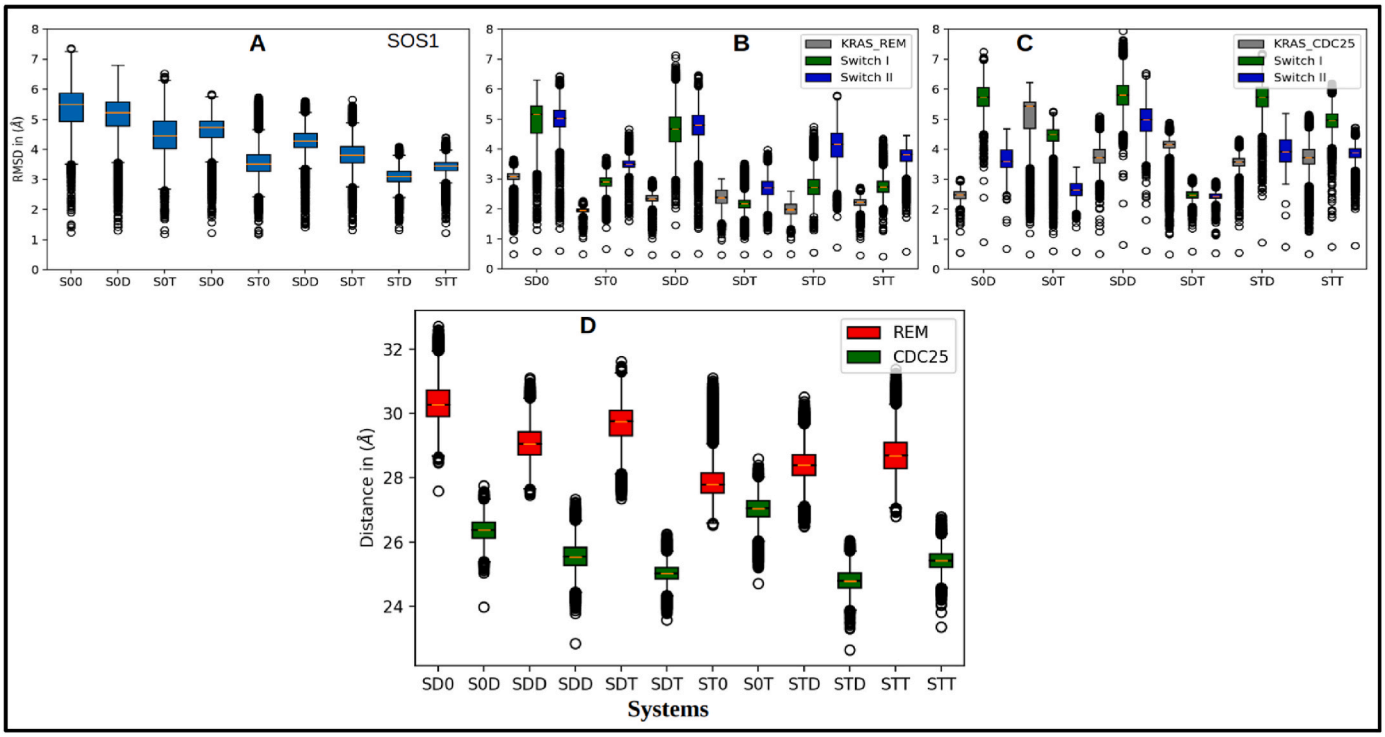

Fig. 3. (A): Distribution of the RMSD values for the SOS1; **(B)** Distribution of the RMSD values for the $KRAS^{REM}$ site systems and switch-I, II of $KRAS^{REM}$; **(C)** Distribution of the RMSD values for the $KRAS^{CDC25}$ systems and switch-I, II of $KRAS^{CDC25}$ (x-axis: Systems and y-axis: RMSD in Å); **(D)** Center of mass distance (in Å) between REM site of SOS1 and $KRAS^{REM}$ (red Color); CDC25 of SOS1 and $KRAS^{CDC25}$ site (green color); (x-axis: Systems and y-axis: Distance in Å). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

27.5 Å (S0T system).

The distance between the GDP-bound KRAS^{CDC25} and SOS1 was observed to reduce significantly in the presence of GDP/GTP-bound KRAS^{REM}. However, in the presence of GTP-bound KRAS^{REM}, the GDP-bound KRAS^{CDC25} was observed to be closer to the SOS1. The distance between GTP-bound KRAS^{CDC25} and SOS1 was observed to decrease, although the reduction in distance was almost the same in the presence of the GDP/GTP-bound KRAS^{REM}. This indicates that the presence of KRAS^{REM} may help provide a better possibility of KRAS^{CDC25} interacting with SOS1. On comparing the proximity of KRAS^{CDC25} to SOS1 in STD (24.2 Å) and STT (25.8 Å) systems, the former was observed to be closer. These distances imply that the GDP-bound KRAS^{CDC25} may be more favored at CDC25 site of SOS1.

3.3. Binding energetics of KRAS-GDP/GTP

The binding affinity of GDP and GTP with KRAS^{CDC25} was calculated for only the systems which consisted of the latter one (Fig. 4A). The average value of the MMPBSA free energy was calculated using the gmx_MMPBSA tool (Valdés-Tresanco et al., 2021). The details of the free energy components have been given in the Supplementary Table 3. It was observed that the binding of GTP was favored as compared to GDP in the absence of KRAS^{REM}. In all the ternary complexes too similar observations were obtained. Whenever, the GDP-bound KRAS^{REM} was present, the GDP or GTP bind KRAS^{CDC25} with similar free energy values. However, in the presence of GTP-bound KRAS^{REM}, GTP was significantly favored to bind the KRAS^{CDC25} as compared to GDP. This infers that the exchange of GDP to GTP was always favored for the KRAS^{CDC25}. This was clearly demarcated in terms of the difference in the free energy of binding between the systems. This difference between S0D and S0T was around 10 kcal/mol, with S0D showing less stable binding of GDP to KRAS^{CDC25}. Similarly, for the SDT and STT systems this difference was around 13 kcal/mol, with STT showing the strongest binding of GTP to KRAS^{CDC25}.

The predicted free energy of binding between KRAS^{REM}/KRAS^{CDC25} with SOS1 in all the systems was calculated using the PRODIGY web server (Fig. 4B). The 500ns snapshot was considered for performing these calculations. It was observed that the KRAS^{REM} binds strongly to the SOS1 in the GTP-bound state as compared to the GDP-bound state. However, in the case of KRAS^{CDC25}, the binding was observed to be weakened for GTP-bound states (S0T, STT) as compared to GDP-bound states (S0D, STD).

During the SOS1 assisted KRAS activation, in the presence of GTP-bound KRAS^{REM}, GDP-bound KRAS^{CDC25} binds to the catalytic site followed by exchange of GDP with GTP. Fig. 4A depicts similar behavior where higher value of free energy of binding (weak binding) was observed for GDP with KRAS^{CDC25} as compared to that of GTP with KRAS^{CDC25}, in the presence of GTP-bound KRAS^{REM}. This may suggest that when GTP-bound KRAS^{REM} is present, GTP binding to KRAS^{CDC25} would be preferred. Once the GTP is bound to KRAS^{CDC25} it may be released from the catalytic site forming a KRAS free catalytic site to indicate the start of a new cycle. This may be interpreted by observing the free energy of binding between the KRAS^{CDC25} and SOS1 for the systems STD and STT (Fig. 4B). The GDP-bound KRAS^{CDC25} (STD) binds strongly to SOS1 as compared to the GTP-bound KRAS^{CDC25} (STT). This may infer that the binding of GTP-bound KRAS^{CDC25} was weak in the STT system, indicating its release from the SOS1 catalytic site. A schematic representation of this activation cycle has been given in Fig. 4C.

3.4. Activation specific crucial distances

The distance and energetics analysis shows that the presence of GDP/GTP bound KRAS^{REM} may have led to conformational changes that differentially regulates the KRAS^{CDC25}. To understand the differential regulation the structural changes in KRAS at both sites in presence of different nucleotides have been studied. Researchers have previously suggested that the KRAS is known to exist in two forms namely, state 1 and state 2. State 1 may be considered as an open-inactive conformation where the switch-I is in an open conformation and state 2 as closed-active where switch-I is in a closed conformation (Spoerner et al.,

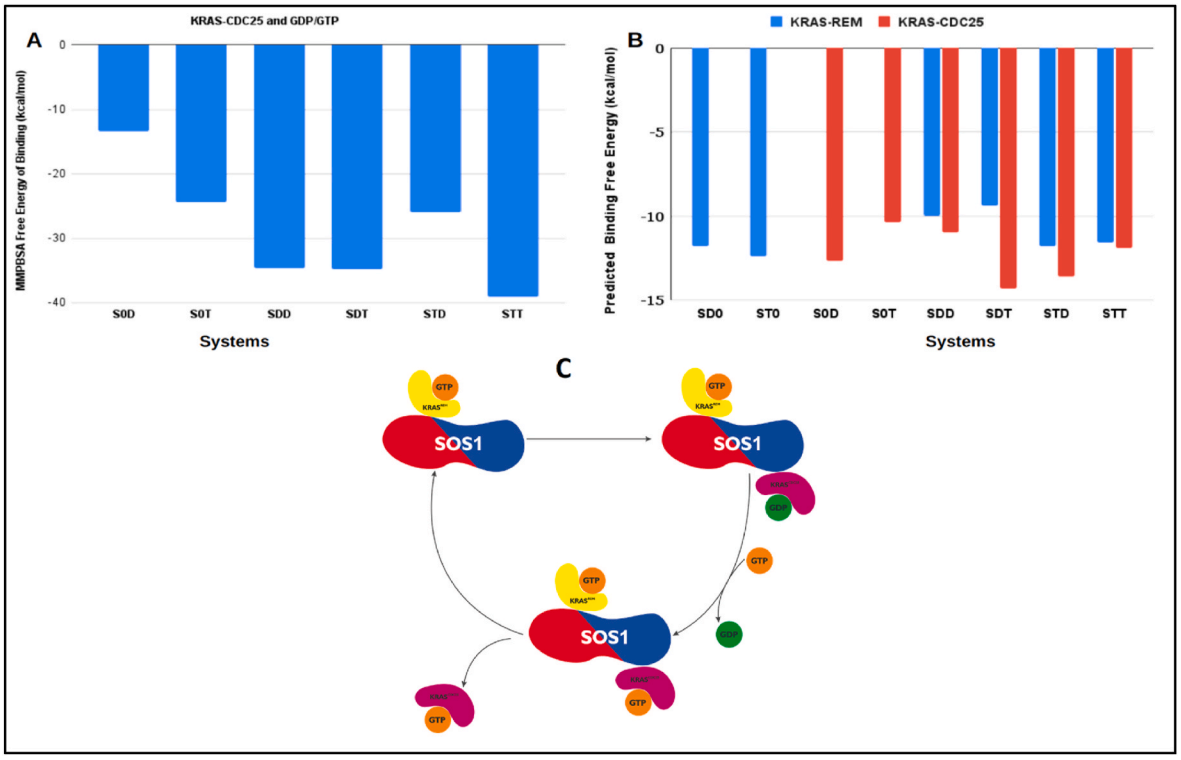

Fig. 4. (A) Average free energy of binding between KRAS and GDP/GTP on either of the binding sites of SOS1; **(B)** Predicted free energy of binding between KRAS^{REM}/KRAS^{CDC25} with SOS1 in all the systems; **(C)** Schematic representation of the activation cycle of KRAS at the catalytic site (STO→STD→STT).

2001). These states have specific conformational signatures, defined based on the distances between KRAS residues and the nucleotides bound to it. These distances are known to be between Y32 and GDP/GTP, T35 and GDP/GTP, G60 and GDP/GTP. The distances of the three KRAS residues namely Y32, T35 and G60 with GDP are known to be greater as compared to their distances with GTP, suggesting state 1 and state 2, respectively (Sharma et al., 2020), (Sharma et al., 2014). In order to observe the variation in these distances, they were calculated for all the nine systems which have been shown in Fig. 5A–C. The distance between Y32 and GDP/GTP at the $KRAS^{REM}$ site was observed to lie in the range of 5–15 Å (Fig. 5A). The maximum distance of ~15 Å was observed for SDD and STT systems. The minimum distance of ~5 Å was observed for the system SD0 and SDT. The distance between Y32 and GDP/GTP at the $KRAS^{CDC25}$ site was also observed to lie in the same range of 5–15 Å (Fig. 5A). The S0D and STD systems were observed to have a maximum distance of ~15 Å. The SDT and STT systems showed a minimum distance of ~5 Å.

The distance between T35 and GDP/GTP at the $KRAS^{REM}$ site and $KRAS^{CDC25}$ site have been calculated and shown in Fig. 5B. The distance between T35 and GDP/GTP at the $KRAS^{REM}$ site was observed to lie in the range of 10–25 Å. The STT system was observed to have a maximum distance of ~25 Å. The system ST0 showed a minimum distance of 9 Å. The distance between T35 and GDP/GTP at the $KRAS^{CDC25}$ site was observed to lie in the range of 6–18 Å. The STD system was observed to have maximum distance ranging around 18 Å. A minimum distance of 7 Å was observed for S0D. The distance between G60 and GDP/GTP at the $KRAS^{REM}$ site and $KRAS^{CDC25}$ site has been given in Fig. 5C. The distance between G60 and GDP/GTP at the $KRAS^{REM}$ site lies in the range of 10–25 Å. The STT system was observed to have a maximum distance of around 25 Å. The minimum distance of 10 Å was observed for the SDT system. The distance between G60 and GDP/GTP at the $KRAS^{CDC25}$ site was observed to lie in the range of 6–17 Å. The SDD and STD systems were observed to have a maximum distance of around 17 Å. The system STT showed a minimum distance of around 5 Å.

The distance calculation between these residues suggest that presence of GTP bound $KRAS^{REM}$ tends to affect the $KRAS^{CDC25}$ to a larger extent as compared to the GDP bound $KRAS^{REM}$. Also, the presence of the GTP bound $KRAS^{CDC25}$ tends to remain mostly in the intermediate state. However, the presence of the GDP bound $KRAS^{REM}$ tends to stabilize the GTP bound $KRAS^{CDC25}$ in the active closed state. One interesting observation was that binding of nucleotide free $KRAS^{REM}$ also tends to stabilize the GTP/GDP bound $KRAS^{CDC25}$ in the active closed state (state 2).

3.5. Region-specific interactions between KRAS and SOS1

The interaction analysis was carried out between the $KRAS^{REM}$ and SOS1 and also between the $KRAS^{CDC25}$ and SOS1 for entire trajectories. The GetContacts tool was used for calculation of contacts. Interactions with occupancy more than 50% throughout simulations have been reported (https://getcontacts.github.io/). Interactions were calculated for all the above-mentioned eight systems as depicted in the flareplots (Supplementary Figure S1 A-F).

The effector binding regions namely P-loop, switch-I and switch-II form the important functional components of KRAS and are known to be involved in interactions with various effector proteins. These regions of KRAS are also known to undergo conformational changes during the activation process of KRAS. Thus, understanding the interaction of KRAS with SOS1 will help in identifying important regions of KRAS which may be playing an important role in the activation mechanism of KRAS. Hence, various interactions between KRAS at both sites and SOS1 were calculated. Fig. 5 (D, E) shows the number of interactions formed between the different regions on KRAS and SOS1 for all the systems. These interactions have been depicted individually in Supplementary Fig. S1 A-F. Fig. 5D depicts the interactions formed between $KRAS^{REM}$ and SOS1. It was observed that the P-loop, beta4, alpha3 and alpha4 regions of $KRAS^{REM}$ had no interactions with the SOS1. It was observed that systems SDD and SDT showed lesser number of interactions with switch-

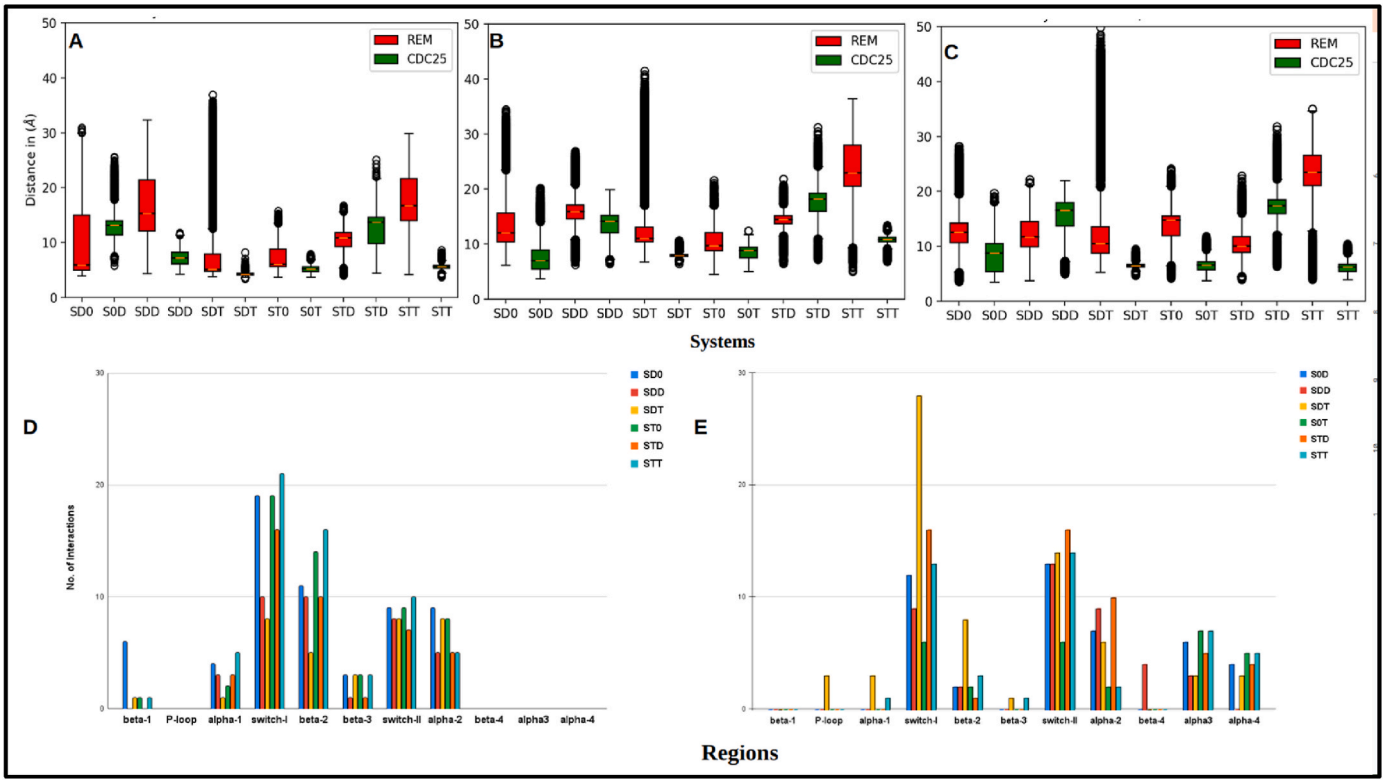

Fig. 5. Distribution of the distance between **(A)** Y32; **(B)** T35 and; **(C)** G60 of KRAS with GDP/GTP (x-axis: Systems and y-axis: Distance in Å). **(D)** Interactions between the crucial regions of **(D)** $KRAS^{REM}$ with SOS1; **(E)** $KRAS^{CDC25}$ with SOS1.

I region as compared to STT. The system STT showed maximum interactions with the switch-I region of KRAS^{REM} (Fig. 5D). The P-loop region of KRAS^{CDC25} was observed to interact with SOS1 only in case of the SDT system (Fig. 5E). The switch-I region of KRAS^{CDC25}, formed maximum interactions with SOS1 in the SDT system. These interactions decreased in the systems STD, STT, S0D, SDD and S0T. For interactions in the switch-II region, system S0T had minimum interactions followed by S0D, SDD, SDT, STT and STD showed maximum interactions. For the interactions in the alpha3 region, SDD and SDT showed minimum interactions, followed by STD, S0D and STT showed maximum interactions. The switch-I of KRAS^{REM} formed a comparatively more number of interactions with SOS1 in the GTP bound state. Besides these regions it was observed that SOS1 mostly showed interaction with beta1, alpha1, beta2, beta3 and alpha2 regions of KRAS^{REM} and alpha3, beta2 and alpha2 regions of KRAS^{CDC25}. It was observed that the presence of GTP at KRAS^{CDC25} tends to increase the interactions of SOS1 with alpha3 regions. It was also observed that the SOS1 interactions were higher in number with the beta2 region of the KRAS^{REM} when bound to GTP. These suggest that the beta2 region of GTP bound KRAS^{REM} may guide the activation through the alpha3 region of KRAS^{CDC25}. The role of the alpha3 region in the activation mechanism of the KRAS has been suggested in some of the previous studies (Haza et al., 2021), (Vatansever et al., 2020). Crucial residue at the end of the alpha3 region namely, K104 is known to be involved in the allosteric network of KRAS upon its acetylation (Yang et al., 2023).

3.6. Dominant structural dynamics

PCA was performed for the SOS1 protein from all nine systems for comparing and characterizing the dominant molecular motions, which may give more objective information to report the overall conformational changes in the system. All the heavy atoms of the SOS1 protein were considered for calculation of eigenvalues and eigenvectors. The decomposition of these eigenvectors was done to obtain the principal component.

3.7. Structural dynamics of SOS1

The distribution plot for conformers along the principal components (PC) 1 & 2 has been given in Supplementary Fig. S2 and RMSF of PC1 of SOS1 region has been given in Fig. 6. It was observed that for most of the systems, the population was spread along both PC's, however a single dominant cluster was observed (Supplementary Fig. S2). Whenever a PC1-PC2 value pair showed the presence of more than 2500 conformations, it was considered to be a cluster. Based on this criterion, only the systems S0T and SDD showed two clusters. However, few of the other systems show the presence of additional clusters which consist of a number of conformers in the range of 2000–2500. In the case of system STT the distribution of population was constrained. In order to see which of the regions showed maximum deviation resulting in these dense clusters, the root mean square fluctuations (RMSF) were calculated along the principal components (PC's). It was observed that for all the systems except the system S00 and system STD, the residue range 661–670 showed fluctuation of around 10 Å. The region 654 to 676 of the REM domain is known to be disordered in the experimental structure (He et al., 2022), (Liao et al., 2018), (Burns et al., 2014). Besides this region, the fluctuating regions differ in different systems. For the system S00, the N-terminal region showed a slightly higher RMSF value of around 5 Å compared to the rest of the protein. For the system S0D the residue range 730–770 showed higher RMSF. The residue range 592–600 showed higher RMSF values in the system SD0. In the system S0T the residue range from 950 to 965 had higher RMSF values. The system SDD showed fluctuation for residue range 723–728, 750–758 and 810–820. The system STD showed fluctuation in multiple regions viz. residue range 730–740, 790–820 and 900–950. The system ST0, SDT and STT didn't show fluctuation in protein except the residue range 661–670. RMSF calculations along PC1 also suggest that only in the case of STD systems, the SOS1 residues within the range 661–670, were observed to be stable as compared to other systems. It was observed that this region which belongs to the allosteric site (REM) made no interactions with the KRAS^{REM} in any of the systems that had KRAS^{REM} present. However, on observing the RMSF of these residues along PC2 for these residues in the STD system, it was observed to fluctuate similarly as the other eight systems did along PC1 (Supplementary Fig. S3).

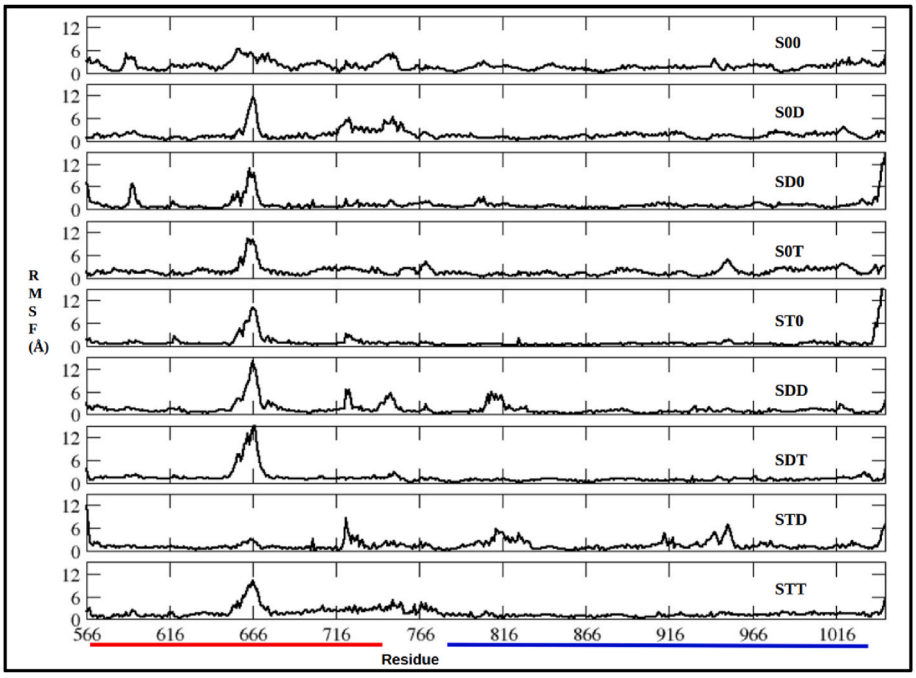

Fig. 6. RMSF of PC1 of SOS1 region, where residues 567–741 represents REM site of SOS1 (red line) and residues 780–1019 represents CDC25 site of SOS1 (blue line). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.8. Structural dynamics of KRAS^{REM}

The dominant motions in individual KRAS protein at either site were identified using PCA. Supplementary Fig. S4 shows the PCA distribution for KRAS^{REM}. It was observed that the KRAS conformers were mostly confined as a single dominant cluster in all the systems. The fluctuations for the residues of KRAS^{REM} systems have been given in Fig. 7. In the case of the SD0 system, the fluctuations were observed only in the switch-I and switch-II region with an average value of 4 Å. In the case of SDD system fluctuations were seen in the entire KRAS protein with an average value of 4.5 Å. For the SDT system major fluctuations were observed in the region connecting alpha3 and alpha4 along with switch-I and switch-II. Similar observations were observed in ST0 as that of SDT system, except in ST0 fluctuations were also observed in the alpha4 region. In the STD system, the switch-II region showed major fluctuation with a value of around 4.5 Å. For the STT system, the switch-I and switch-II region showed major fluctuations. It was observed that irrespective of the nucleotide bound to KRAS, in all the systems major fluctuations were observed in the switch-I and switch-II regions. Besides these two regions, in some of the systems the fluctuations were observed in alpha3, alpha4 and the region connecting alpha3 and alpha4. This suggests that switch-I, switch-II, alpha3, alpha4 and the region between alpha3 and alpha4 may be playing an important role in communicating allosteric effects to KRAS at the active site through SOS1 protein.

3.9. Structural dynamics of KRAS^{CDC25}

The distribution of conformers along PC1 and PC2 shows that the motions in KRAS^{CDC25} systems showed more than one populated cluster Supplementary Fig. S5). RMSF for the residues of KRAS at the catalytic site were also calculated (Fig. 8). Similar to KRAS^{REM}, the KRAS^{CDC25} showed similar behavior except for the system SDT. For the S0D system, switch-I showed major fluctuation (around 4 Å) followed by switch-II (around 3 Å), alpha3 (around 3 Å) and alpha4 (around 2.5 Å). In the case of the SDD system, switch-II showed major fluctuation of around 4 Å and switch-I around 3 Å. The KRAS^{CDC25} for the SDT system was most stable with an average fluctuation of 3 Å for all the regions. The S0T showed fluctuation in switch-I, switch-II and C-terminal regions. The STD showed fluctuation for switch-II and alpha4 region with average value of around 3 Å. For STT, fluctuations were observed in switch-I, switch-II and alpha4 regions with an average value of around 3 Å. Thus, the presence of GDP at the allosteric site and GTP at the catalytic site had the most stable system.

In order to see which state (active and inactive) of KRAS dominates the PCA clusters the distance between the important residues known to play a role during the activation and release of nucleotides from KRAS were calculated for the representative structures of each cluster. The distance between Y32 and GDP/GTP, T35 and GDP/GTP and G60 and GDP/GTP were calculated and corresponding distance values have been depicted in Supplementary Table 4. For the S0D system, the KRAS^{CDC25} was mostly in state 1. For the S0T system, the KRAS^{CDC25} was mostly in an intermediate state. The KRAS^{REM} in the SD0 system showed conformations of the intermediate state. The KRAS^{REM} and KRAS^{CDC25} in the SDD system were in state 1. The KRAS^{REM} in the SDT system was in state 1 while the KRAS^{CDC25} system showed confirmation of the intermediate state. The KRAS^{REM} in the ST0 system showed the state 1 conformation. The KRAS^{REM} and KRAS^{CDC25} in the STD system showed conformations that were in the state 1. The KRAS^{REM} in the STT system showed intermediate state conformation and KRAS^{CDC25} showed active closed state conformation.

Among all the systems, it was found in system STD, the KRAS^{CDC25} showed open-inactive state conformation with maximum distance between crucial residue and nucleotide essential for holding the nucleotide. Thus it may suggest that the probability of release of nucleotide from KRAS^{CDC25} may be maximum in STD system and also may promote an increased rate of the nucleotide exchange. On the other hand, in the STT system, the distance between crucial residues of KRAS^{CDC25} and GTP is minimum, representing closed active state structure, thereby making the release of the nucleotide GTP from KRAS^{CDC25} rarely possible.

3.10. Interactions guiding inter-state transition in KRAS

The interactions at the interface of KRAS-SOS1 that may be

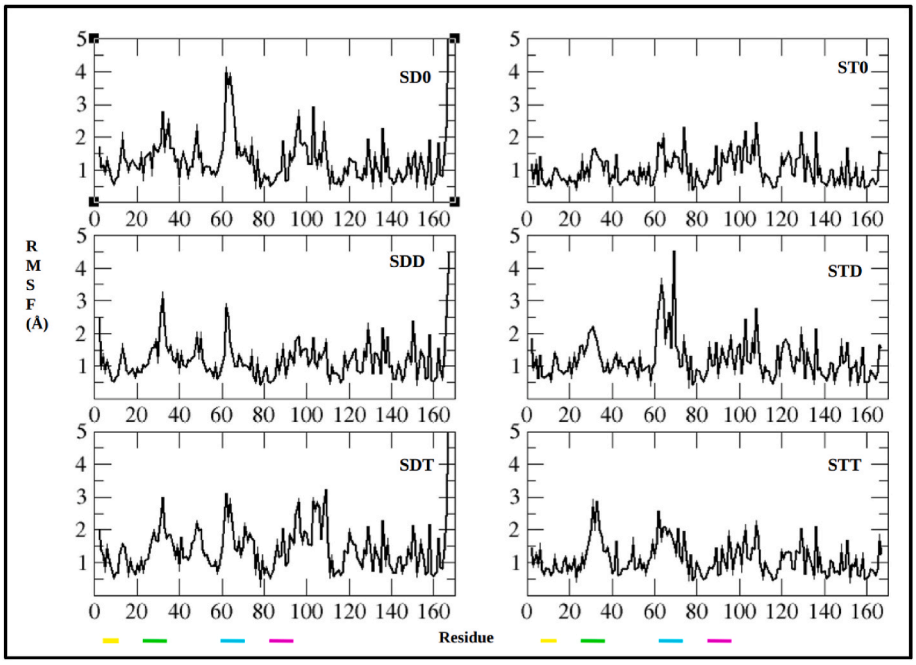

Fig. 7. RMSF of KRAS^{REM} site residues along PC1 (P-loop region: residue number 10–14 (yellow color), switch-I: residue number 30–40 (green color), switch-II: residue number 58–72 (cyan color), alpha3: residue number 87–104 (magenta color), alpha4: residue number 127–137). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

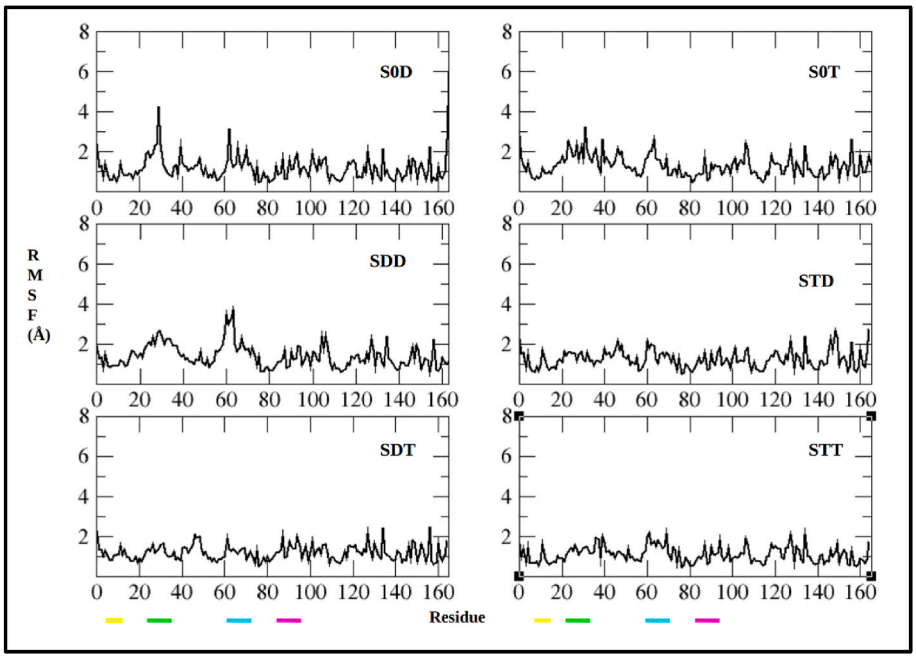

Fig. 8. RMSF of KRAS^{CDC25} site along PC1 (P-loop region: residue number 10–14 (yellow color), switch-I: residue number 30–40 (green color), switch-II: residue number 58–72 (cyan color), alpha3: residue number 87–104 (magenta color), alpha4: residue number 127–137). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

responsible for transition of KRAS from one state to another were calculated for all the systems and has been given in Supplementary Fig. S1 A-F and Supplementary Table 5. Some of the interactions formed by KRAS^{REM} and KRAS^{CDC25} with SOS1 have been color coded and shown in Fig. 9 A and 9 B, respectively. The KRAS^{REM} residues were observed to form more interactions with residues of SOS1 as compared to KRAS^{CDC25} residues. The interacting residue pairs discussed further, to be considered as KRAS^{REM} residue - SOS1 residue (Fig. 9A). Q25-N976

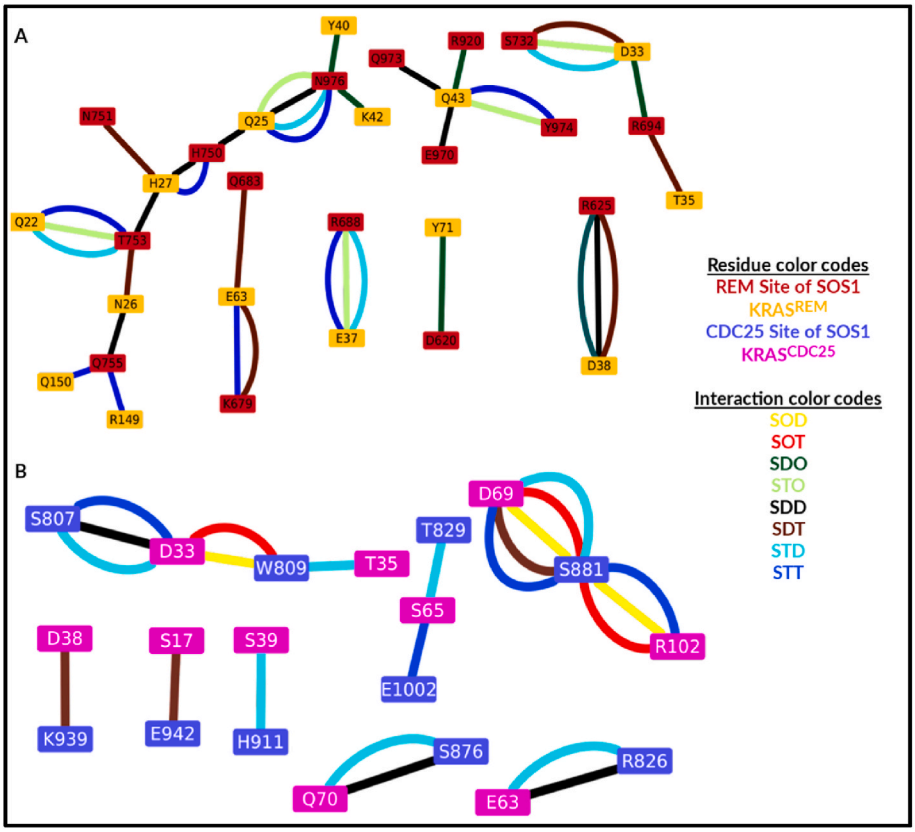

Fig. 9. Interactions between residues of SOS1 and **(A)** KRAS^{REM} and **(B)** KRAS^{CDC25} for all the systems. Color representations have been given in the figure. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

was observed to be stable in most simulation systems namely STO, SDD, STD and STT. Q22-T753 and E37-R688 were found to be stable in three simulation systems namely STT, STO and STD. Additionally, D33-S732 and D38-R625 were observed to be stable in three simulation systems, out of which SDT was the common one. D33-S732 was also observed in STO and STD and D38-R625 in SD0 and SDD. The most number of interactions, eight were formed by the systems SDD and STT. Q25-N976 and H27–H750 were ones that were observed in both. The interacting residue pairs discussed further, to be considered as $KRAS^{CDC25}$ residue-SOS1 residue (Fig. 9 B). D69-S881 was observed to be stable in five simulation systems namely STT, SDT, SOD, SOT and STD. The GTP-bound $KRAS^{REM}$ systems, namely, STO and STD were observed to have four interactions in common. These interactions were Q22-T753, Q25-N976, D33-S732 and E37-R688. It may be inferred that these interactions may play a crucial role in employing the GDP-bound KRAS at the CDC25 (catalytic) site of SOS1 (Liao et al., 2018).

For the SD0 system the switch-I and switch-II residues of $KRAS^{REM}$ were mainly engaged in interaction with SOS1 protein. The residue D33, D38 and Y40 of switch-I region formed interactions with R694, R625 and N975 residues of SOS1 respectively. The residues S65 and Y71 of switch-II region formed interactions with R688 and D620 respectively. Besides these, the residue of the linker region connecting the switch-I and switch-II residues Y40 and K42 formed interactions with N976 and the R41 and Q43 formed interactions with E966 and R920 respectively. In the case of the ST0 system, the majority of interactions between SOS1 and $KRAS^{REM}$ belong to only the switch-I region of $KRAS^{REM}$. The D33 and E37 residues of KRAS formed interactions with S732 and R688 residues of SOS1 respectively. The linker region residues R41 and Q43 formed interactions with N622 and Y974. It was observed that the presence of GTP instead of GDP favored more interactions between switch-I and switch-II regions of KRAS with SOS1.

In the SDD system, for $KRAS^{REM}$ the switch-I residues D33 and D38 formed interactions with R694 and R625 respectively. For $KRAS^{REM}$, in the STD system, the switch-I of $KRAS^{REM}$ showed interaction with SOS1, the residues D33 showed interactions with R694 and S732. Along with switch-I and switch-II, the N-terminal residue Q22 formed an interaction with T753. The presence of GTP bound $KRAS^{CDC25}$ (SDT) enhanced the interaction between KRAS at the catalytic site and SOS1. The switch-I residues D33, T35 and D38 formed interactions with S732, R694 and R625 respectively. The switch-II residue, E63 formed interactions with K679 and Q683. For the system STT, where GTP bound KRAS is present at both catalytic and allosteric sites, it was observed that switch-I residues of $KRAS^{REM}$ viz. D33 and E37 showed interactions with R688. The switch-II residues of $KRAS^{REM}$ E63 and Y64 showed interactions with K679 and Q683 respectively. The C-terminal residue of $KRAS^{REM}$ R149 showed interaction with Q755.

For the system S0D, the switch-I residue of $KRAS^{CDC25}$ D33 formed an interaction with W809 of SOS1. Whereas, the switch-II residues of $KRAS^{CDC25}$, Q61 and D69 formed interactions with K814 and S881 respectively. The linker region residue R73 forms an interaction with S881. The alpha3 residues H95 and R102 formed interactions with K1003 and S881 respectively. D105 forms an interaction with R1019. For the system S0T, the switch-I residues of $KRAS^{CDC25}$, namely D33, E37 and D38 formed interactions with S807, H911 and K939 of SOS1 respectively. The switch-II residues of $KRAS^{CDC25}$ E63, S65, R68 and D69 formed interactions with R826, E1002, E1002 and S881 respectively. The alpha3 residue R102 formed an interaction with S881. The D105 forms an interaction with R1019. For the system SDD, linker region residue Q25 forms an interaction with T940 and the switch-I residues D33, D38 form interactions with S807, K913 respectively. The linker region residue R41 forms an interaction with D910. The switch-II residues E63, Q70, Q70 and Y71 form interactions with R826, S876, Y912 and Y912 respectively. The residue R73 forms interactions with N879 and S881. For the system STD, switch-I residues D33, D33, T35 and S39 form interactions with S807, W809, W809 and H911 respectively. The linker region residues R41 and D54 form interactions with D910 and H911 respectively. The switch-II residues E62, S65, D69 form interactions with R826, T829, and S881 respectively. The switch-II residue Q70 forms interactions with S876, H905 and Y912. In the system SDT, the P-loop residue S17 forms an interaction with E942. The switch-I residues E31, E37, D38 and S39 form interactions with K960, Y912, K939 and H911 respectively. The switch-II residues E63, D69 and Q70 form interactions with R826, S881 and H905 respectively. The linker region residue R73 forms an interaction with N879. The alpha3 region residue R102 forms an interaction with S881. The D105 forms an interaction with R1019. In the system STT, switch-I residues of $KRAS^{CDC25}$ D33, E37 and D38 form interactions with S807, Y912 and K939 respectively. The switch-II residue Q61, E63, S65, R68 and D69 form interactions with W809, K814, E1002, E1002 and S881 respectively. The alpha3 residues D92, R102 and R102 form interactions with K1003, S881 and T1006 respectively. D105 also forms an interaction with R1019.

4. Discussion

The experimental and theoretical studies have reported that the presence of the different nucleotide bound KRAS at the allosteric site (REM site of SOS1) influences the activation at the catalytic site (CDC25 site of SOS1) (Liao et al., 2018). Hence, in the current study to understand the SOS1 mediated activation of KRAS, in all nine systems with different combinations of nucleotide bound KRAS at REM and CDC25 site of SOS1 have been studied using MD simulations. The structural changes in $KRAS^{REM}$-SOS1-$KRAS^{CDC25}$, $KRAS^{REM}$-SOS1 and SOS1-$KRAS^{CDC25}$ systems which may guide the activation of KRAS were explored extensively. The work by Liao TJ et al. had discussed three different activation cycles viz. fast, slow and rare activation cycles, depending on the presence of GDP/GTP bound to the $KRAS^{REM}$ site (Liao et al., 2018). The fast cycle was represented by the presence of GTP at $KRAS^{REM}$ leading to fast transition of nucleotides at the catalytic site i.e. from ST0→STD→STT. The slow cycle was represented by the presence of GDP at $KRAS^{REM}$ leading to slow transition of nucleotides at activation site i.e. SD0→SDD→SDT. The rare cycle was represented by the absence of nucleotide at $KRAS^{REM}$ leading to a very rare transition of nucleotide at activation site i.e. S0D→S0T. The free energy analysis in the current study depicts a stronger binding between GTP and $KRAS^{CDC25}$ as compared to the GDP and $KRAS^{CDC25}$ indicating probable GDP to GTP transition. The free energy of binding calculated between the nucleotides and $KRAS^{CDC25}$ indicated that most favored GDP to GTP exchange occurred in STD→STT (fast) followed by SOD→SOT (rare). The SDD→SDT (slow) transition was the least favored. These observations may infer that fast and rare activation cycles were more favored as compared to the slow cycle for activation of KRAS at the catalytic site.

In most of the activation studies on KRAS, it was reported that switch-I and switch-II regions show considerable dynamics (Chen et al., 2013), (He et al., 2022), (Liao et al., 2018). It is known that these switch regions interact with different effector proteins and are also involved in activation of KRAS protein. The study by Liao TJ et al. showed that these switch regions tend to form hydrophobic, salt-bridge, H-bond and polar interactions with SOS1 protein (Liao et al., 2018). Xuan He et al. in their study showed that the switch-I residues Y32, D30 and switch-II residues Q61 and E63 residues of KRAS are engaged in the interactions with N547, K566, K814 and R826 residues of SOS1 respectively (He et al., 2022). Burns MC et al. showed that along with switch-I and switch-II regions the R73 forms interaction with N879 (Burns et al., 2014). Hillig RC et al. have also shown the role of R73 from KRAS and N879 from SOS1 to form interactions with inhibitors by obstructing the reloading of KRAS with GTP, resulting in significant anti-proliferative effects (Hillig et al., 2019). The switch-I residues Y32 and T35 and switch-II residues G60 and G61 have been reported to play an important role in GTP hydrolysis (Sharma et al., 2014).

Similar to above studies, in the current work, it was observed that the residues from switch-I and switch-II regions of the $KRAS^{REM}$ site formed

interactions with SOS1. For the switch-I region, the interactions were D33-R688, D33-R694, D33-S732, T35-R694, E37-R688, D38-R625 and Y40–N976. For the switch-II region, the interactions were E63-K679, E63-Q683, Y64-Q683, S65-R688 and Y71-D620. The switch-I region of $KRAS^{CDC25}$ showed interactions with SOS1, the KRAS-SOS1 residue pairs involved were D33-S807, D33-W809, E37-H911 and D38-K939. Similarly, for the switch-II region were Q61-K814, E63-R826, S65-E1002 and D69-S881. Also, the alpha3 region residues of the $KRAS^{CDC25}$ were involved in the following interactions with SOS1, H95–K1003 and R102-S881. R149 formed an interaction with Q755. Further, it was observed that the distance between the $KRAS^{REM}$ and SOS1 showed more variation among different systems as compared to distance between S0S1 and $KRAS^{CDC25}$. The interaction analysis suggests that the residues from switch-I, switch-II of $KRAS^{REM}$ may be involved in initiation of the allosteric signals. Through the SOS1 this allosteric signals may be transmitted to switch-I, switch-II and alpha 3 of $KRAS^{CDC25}$. The role of switch-I and switch-II in activation of KRAS has been discussed in many studies (He et al., 2022), (Liao et al., 2018), (Yang et al., 2023). K104 which is a part of the alpha3 region of KRAS is known to be involved in the allosteric activation network on its acetylation (Yang et al., 2023). The principal component analysis in the current study showed high fluctuation in the region with residue range 661–670 of SOS1 and the residues flanking this region were observed to interact with the switch-I of KRAS. Additionally, the nearby residues to this region are also known to be critical for allosteric interaction with KRAS. The residue E37 of the switch-I region of KRAS is known to interact with R688, N691 and H695 of SOS1, thereby forming the interacting interface between $KRAS^{REM}$ and SOS1 (Liao et al., 2018). The switch-I region is known to interact with switch-II and GTP, suggesting that GTP hydrolysis may trigger structural changes in the switch-I. The regions of SOS1, lying within the range 800–820 (SDD/STD), 911–914 (STD) and 940–953 (STD), known to interact with $KRAS^{CDC25}$ showed structural variation in the simulations. This suggests that these regions may play a critical role in mitigating the differential allosteric effect in activating KRAS at the catalytic site. Especially, the region 940–953 of SOS1 forms a part of the helix hairpin motif which is known to impart allosteric effects from REM site to CDC25 site (Freedman et al., 2006). Q43 of $KRAS^{CDC25}$ is known to interact with the sidechain hydroxyl group of Y974 (Liao et al., 2018). This residue of $KRAS^{CDC25}$ was observed to form hydrogen bond interaction with backbone carbonyl groups of L919 and I922 of the CDC25 site of SOS1. Additionally along with the residues of switch-I and switch-II regions, it was observed that the alpha3, beta1, alpha1, beta2, beta3 and alpha2 regions of KRAS were involved in forming interaction with SOS1. The current study highlights important regions and residues that may play a role in transmitting allosteric signals from REM site of SOS1 to CDC25 site and suggest their probable role in activation of KRAS. There are studies known where rescue strategies have been designed to halt KRAS activation by targeting the KRAS-SOS1 interface (Ketcham et al., 2022), (Hillig et al., 2019), (Fang et al., 2018). Hence, the information on the interacting residues of KRAS and SOS1 at the catalytic site obtained through the reported simulations may prove to be useful in abrogating the SOS1-mediated KRAS activation. This information may be useful for designing a potential therapeutic approach to inhibit the activation of KRAS catalyzed by the GEF: SOS1.

5. Conclusion

In the current study an attempt has been made to understand the different events during the SOS1 mediated KRAS activation. Molecular dynamics simulations were performed for all the nine systems viz. dimer systems consisting of KRAS-SOS1 and ternary complex consisting of KRAS-SOS1-KRAS in the presence GTP/GDP. It was observed that switch-I, switch-II and alpha3 regions of KRAS were found to be playing an important role in the activation process. The $KRAS^{REM}$ residues from switch-I regions (D33, T35, E37, D38 and Y40) and from switch-II regions (E63, Y64, S65 and Y71) were found to be forming interactions with D620, R625, K679, Q683 R688, R694, S732 residues of SOS1. Similarly, switch-I residues of $KRAS^{CDC25}$ namely D33, D37 and D38 were involved in interaction with S807, W809, H911 and K939 of SOS1 and the switch-II residues namely E63, Q61, S65 and D69 formed interactions with R826, K814, E1002 and S881 of SOS1 respectively. These interactions may be important for allosteric signaling, as the residues are located at the interface between KRAS and SOS1. The SOS1 regions with residue range 661–670, 723 to 758, 800 to 825, 911 to 914 and 940 to 953 showed maximum rearrangement during the course of simulations and may suggest their role in mitigation of allosteric signals. The study showed that GTP was more favored over GDP at the $KRAS^{CDC25}$ site. Further to this, overall structural variation observed suggests that both fast (ST0→STD→STT) and slow (SD0→SDD→SDT) activation cycles were explored in the $KRAS^{REM}$-SOS1-$KRAS^{CDC25}$ ternary complex simulations. The systems of fast activation cycle (ST0→STD→STT) had common interactions namely Q22-T753, Q25-N976 and E37-R688. The systems of slow activation cycle (SD0→SDD→SDT) had common interaction namely D38-R625. The systems of rare activation cycle (S0D→S0T) had common interaction namely D33-S807, D69-S881 and R102-S881. The study reports important interaction with their probable role in activations mechanism and may hold important insight for further understanding of SOS1 mediated KRAS activation mechanism.

Author statement

Kirti Bhadhadhara, Vinod Jani and Shruti Koulgi conceptualized the idea. Kirti Bhadhadhara conceived the research work, did analyses of data. Kirti Bhadhadhara, Vinod Jani and Shruti Koulgi wrote the manuscript. Dr. Uddhavesh Sonavane and Dr. Rajendra Joshi edited and reviewed the final manuscript.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the Ministry of Electronics and Information Technology (MeitY), Government of India and the National Supercomputing Mission (NSM) as the funding sources. This work was performed using the supercomputing facility: The Bioinformatics Resources and Applications Facility (BRAF) at C-DAC Pune, India.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crstbi.2023.100115.

References

Bandaru, P., Kondo, Y., Kuriyan, J., 2019. The interdependent activation of Son-of-Sevenless and Ras. Cold Spring Harbor perspectives in medicine 9 (2), a031534.
Berman, H., Henrick, K., Nakamura, H., 2003. Announcing the worldwide protein data bank. Nat. Struct. Mol. Biol. 10 (12), 980.
Burns, M.C., Sun, Q., Daniels, R.N., Kennedy, J.P., Phan, J., Olejniczak, E.T., Lee, T., Waterson, A.G., Rossanese, O.W., Fesik, S.W., 2014. Approach for targeting Ras with small molecules that activate SOS-mediated nucleotide exchange. Proc. Natl. Acad. Sci. USA 111 (9), 3401–3406.
Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I., Brozell, S.R., Cerutti, D.S., Cheatham III, T.E., Cruzeiro, V.W., Darden, T.A., Duke, R.E., Giambasu, G., 2021. Amber 2021. University of California, San Francisco.

Chen, C.C., Er, T.K., Liu, Y.Y., Hwang, J.K., Barrio, M.J., Rodrigo, M., Garcia-Toro, E., Herreros-Villanueva, M., 2013. Computational analysis of KRAS mutations: implications for different effects on the KRAS p. G12D and p. G13D mutations. PLoS One 8 (2), e55793.
Chen, J., Wang, L., Wang, W., Sun, H., Pang, L., Bao, H., 2021. Conformational transformation of switch domains in GDP/K-Ras induced by G13 mutants: an investigation through Gaussian accelerated molecular dynamics simulations and principal component analysis. Comput. Biol. Med. 135, 104639.
Chong, L.T., Pitera, J.W., Swope, W.C., Pande, V.S., 2009. Comparison of computational approaches for predicting the effects of missense mutations on p53 function. J. Mol. Graph. Model. 27 (8), 978–982.
Fang, Z., Marshall, C.B., Nishikawa, T., Gossert, A.D., Jansen, J.M., Jahnke, W., Ikura, M., 2018. Inhibition of K-RAS4B by a unique mechanism of action: stabilizing membrane-dependent occlusion of the effector-binding site. Cell Chem. Biol. 25 (11), 1327–1336.
Fernandes, H.S., Sousa, S.F., Cerqueira, N.M., 2019. VMD store-A VMD plugin to browse, discover, and install VMD extensions. J. Chem. Inf. Model. 59 (11), 4519–4523.
Freedman, T.S., Sondermann, H., Friedland, G.D., Kortemme, T., Bar-Sagi, D., Marqusee, S., Kuriyan, J., 2006. A Ras-induced conformational switch in the Ras activator Son of sevenless. Proc. Natl. Acad. Sci. USA 103 (45), 16692–16697.
Gasper, R., Wittinghofer, F., 2020. The Ras switch in structural and historical perspective. Biol. Chem. 401 (1), 143–163.
Gebregiworgis, T., Kano, Y., St-Germain, J., Radulovich, N., Udaskin, M.L., Mentes, A., Huang, R., Poon, B.P., He, W., Valencia-Sama, I., Robinson, C.M., 2021. The Q61H mutation decouples KRAS from upstream regulation and renders cancer cells resistant to SHP2 inhibitors. Nat. Commun. 12 (1), 1–5.
Gimple, R.C., Wang, X., 2019. RAS: striking at the core of the oncogenic circuitry. Front. Oncol. 9, 965.
Gurung, A.B., Bhattacharjee, A., 2015. Significance of Ras signaling in cancer and strategies for its control. In: Journal-Significance of Ras Signaling in Cancer and Strategies for its Control.
Haza, K.Z., Martin, H.L., Rao, A., Turner, A.L., Saunders, S.E., Petersen, B., Tiede, C., Tipping, K., Tang, A.A., Ajayi, M., Taylor, T., 2021. RAS-inhibiting biologics identify and probe druggable pockets including an SII-α3 allosteric site. Nat. Commun. 12 (1), 1–5.
He, X., Du, K., Wang, Y., Fan, J., Li, M., Ni, D., Lu, S., Bian, X., Liu, Y., 2022. Autopromotion of K-Ras4B feedback activation through an SOS-mediated long-range allosteric effect. Front. Mol. Biosci. 9.
Hillig, R.C., Sautier, B., Schroeder, J., Moosmayer, D., Hilpmann, A., Stegmann, C.M., Werbeck, N.D., Briem, H., Boemer, U., Weiske, J., Badock, V., 2019. Discovery of potent SOS1 inhibitors that block RAS activation via disruption of the RAS–SOS1 interaction. Proc. Natl. Acad. Sci. USA 116 (7), 2551–2560.
Hoang, H.M., Umutesi, H.G., Heo, J., 2021. Allosteric autoactivation of SOS and its kinetic mechanism. Small GTPases 12 (1), 44–59.
Iversen, L., Tu, H.L., Lin, W.C., Christensen, S.M., Abel, S.M., Iwig, J., Wu, H.J., Gureasko, J., Rhodes, C., Petit, R.S., Hansen, S.D., 2014. Ras activation by SOS: allosteric regulation by altered fluctuation dynamics. Science 345 (6192), 50–54.
Jeng, H.H., Taylor, L.J., Bar-Sagi, D., 2012. Sos-mediated cross-activation of wild-type Ras by oncogenic Ras is essential for tumorigenesis. Nat. Commun. 3 (1), 1–8.
Kessler, D., Gerlach, D., Kraut, N., McConnell, D.B., 2021. Targeting son of sevenless 1: the pacemaker of KRAS. Curr. Opin. Chem. Biol. 62, 109–118.
Ketcham, J.M., Haling, J., Khare, S., Bowcut, V., Briere, D.M., Burns, A.C., Gunn, R.J., Ivetac, A., Kuehler, J., Kulyk, S., Laguer, J., 2022. Design and discovery of MRTX0902, a potent, selective, brain-penetrant, and orally bioavailable inhibitor of the SOS1: KRAS protein–protein interaction. J. Med. Chem. 65 (14), 9678–9690.
Liao, T.J., Jang, H., Fushman, D., Nussinov, R., 2018. Allosteric KRas4B can modulate SOS1 fast and slow Ras activation cycles. Biophys. J. 115 (4), 629–641.
Link:https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga (accessed 2022-November-10)..
Link:https://getcontacts.github.io/(accessed 2022-December-10)..
Link: https://plasma-gate.weizmann.ac.il/pub/grace/(accessed 2022-December-15)..
Liu, P., Wang, Y., Li, X., 2019. Targeting the untargetable KRAS in cancer therapy. Acta Pharm. Sin. B 9 (5), 871–879.
López, C.A., Agarwal, A., Van, Q.N., Stephen, A.G., Gnanakaran, S., 2021. Unveiling the dynamics of KRAS4b on lipid model membranes. J. Membr. Biol. 254 (2), 201–216.
Lu, S., Jang, H., Nussinov, R., Zhang, J., 2016. The structural basis of oncogenic mutations G12, G13 and Q61 in small GTPase K-Ras4B. Sci. Rep. 6 (1), 1–5.
Lukman, S., Grant, B.J., Gorfe, A.A., Grant, G.H., McCammon, J.A., 2010. The distinct conformational dynamics of K-Ras and H-Ras A59G. PLoS Comput. Biol. 6 (9), e1000922.
Margarit, S.M., Sondermann, H., Hall, B.E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., Kuriyan, J., 2003. Structural evidence for feedback activation by Ras· GTP of the Ras-specific nucleotide exchange factor SOS. Cell 112 (5), 685–695.
Maurer, T., Garrenton, L.S., Oh, A., Pitts, K., Anderson, D.J., Skelton, N.J., Fauber, B.P., Pan, B., Malek, S., Stokoe, D., Ludlam, M.J., 2012. Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. Proc. Natl. Acad. Sci. USA 109 (14), 5299–5304.
Miller III, B.R., McGee Jr., T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E., 2012. MMPBSA.py: an efficient program for end-state free energy calculations. J. Chem. Theor. Comput. 8 (9), 3314–3321.
Moghadamchargari, Z., Shirzadeh, M., Liu, C., Schrecke, S., Packianathan, C., Russell, D. H., Zhao, M., Laganowsky, A., 2021. Molecular assemblies of the catalytic domain of SOS with KRas and oncogenic mutants. Proc. Natl. Acad. Sci. USA 118 (12), e2022403118.
Muñoz-Maldonado, C., Zimmer, Y., Medová, M., 2019. A comparative analysis of individual RAS mutations in cancer biology. Front. Oncol. 9, 1088.
Murugan, A.K., Grieco, M., Tsuchida, N., 2019. RAS mutations in human cancers: roles in precision medicine. InSeminars in cancer biology 59, 23–35. Academic Press.
Páll, S., Zhmurov, A., Bauer, P., Abraham, M., Lundborg, M., Gray, A., Hess, B., Lindahl, E., 2020. Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. J. Chem. Phys. 153 (13), 134110.
Pantsar, T., 2020. The current understanding of KRAS protein structure and dynamics. Comput. Struct. Biotechnol. J. 18, 189–198.
Parker, M.I., Meyer, J.E., Golemis, E.A., Dunbrack Jr., R.L., 2022. Delineating the RAS conformational landscape. Cancer Res. 82 (13), 2485.
Pettersen, E.F., et al., 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25 (13), 1605–1612. https://doi.org/10.1002/jcc.20084.
Saraste, M., Sibbald, P.R., Wittinghofer, A., 1990. The P-loop-a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. 15 (11), 430–434.
Sayyed-Ahmad, A., Prakash, P., Gorfe, A.A., 2017. Distinct dynamics and interaction patterns in H-and K-Ras oncogenic P-loop mutants. Proteins: Struct., Funct., Bioinf. 85 (9), 1618–1632.
Schlagenhauf, P., Adamcova, M., Regep, L., Schaerer, M.T., Rhein, H.G., 2010. The position of mefloquine as a 21st century malaria chemoprophylaxis. Malar. J. 9 (1), 1–5.
Sharma, N., Sonavane, U., Joshi, R., 2014. Probing the wild-type HRas activation mechanism using steered molecular dynamics, understanding the energy barrier and role of water in the activation. Eur. Biophys. J. 43 (2), 81–95.
Sharma, N., Sonavane, U., Joshi, R., 2020. Comparative MD simulations and advanced analytics based studies on wild-type and hot-spot mutant A59G HRas. PLoS One 15 (10), e0234836.
Sondermann, H., Soisson, S.M., Boykevisch, S., Yang, S.S., Bar-Sagi, D., Kuriyan, J., 2004. Structural analysis of autoinhibition in the Ras activator Son of sevenless. Cell 119 (3), 393–405.
Spoerner, M., Herrmann, C., Vetter, I.R., Kalbitzer, H.R., Wittinghofer, A., 2001. Dynamic properties of the Ras switch I region and its importance for binding to effectors. Proc. Natl. Acad. Sci. USA 98 (9), 4944–4949.
Sun, Q., Burke, J.P., Phan, J., Burns, M.C., Olejniczak, E.T., Waterson, A.G., Lee, T., Rossanese, O.W., Fesik, S.W., 2012. Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. Angew. Chem. Int. Ed. 51 (25), 6140–6143.
Uras, I.Z., Moll, H.P., Casanova, E., 2020. Targeting KRAS mutant non-small-cell lung cancer: past, present and future. Int. J. Mol. Sci. 21 (12), 4325.
Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A., Moreno, E., 2021. gmx_MMPBSA: A new tool to perform end-state free energy calculations with GROMACS. J. Chem. Theor. Comput. 17 (10), 6281–6291.
Vatansever, S., Erman, B., Gümüş, Z.H., 2020. Comparative effects of oncogenic mutations G12C, G12V, G13D, and Q61H on local conformations and dynamics of K-Ras. Comput. Struct. Biotechnol. J. 18, 1000–1011.
Venkatakrishnan, A.J., Fonseca, R., Ma, A.K., Hollingsworth, S.A., Chemparathy, A., Hilger, D., Kooistra, A.J., Ahmari, R., Babu, M.M., Kobilka, B.K., Dror, R.O., 2019. Uncovering patterns of atomic interactions in static and dynamic structures of proteins. bioRxiv, 840694.
Vo, U., Vajpai, N., Flavell, L., Bobby, R., Breeze, A.L., Embrey, K.J., Golovanov, A.P., 2016. Monitoring Ras interactions with the nucleotide exchange factor son of sevenless (Sos) using site-specific NMR reporter signals and intrinsic fluorescence. J. Biol. Chem. 291 (4), 1703–1718.
Walker, J.E., Saraste, M., Runswick, M.J., Gay, N.J., 1982. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1 (8), 945–951.
Wang, W.H., Yuan, T., Qian, M.J., Yan, F.J., Yang, L., He, Q.J., Yang, B., Lu, J.J., Zhu, H., 2021. Post-translational modification of KRAS: potential targets for cancer therapy. Acta Pharmacol. Sin. 42 (8), 1201–1211.
Waterhouse, A., et al., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46 (W1), W296–W303. https://doi.org/10.1093/nar/gky427.
Zhao, J., Wang, L., Bao, H., Chen, J., 2021. Cluster analysis on conformational changes of the GDP/KRAS complex induced by A59G and D33E. Chem. Phys. Lett., 138995, 781.
Yang, M.H., Tran, T.H., Hunt, B., Agnor, R., Johnson, C.W., Shui, B., Waybright, T.J., Nowak, J.A., Stephen, A.G., Simanshu, D.K., Haigis, K.M., 2023. Allosteric regulation of switch-II domain controls KRAS oncogenicity. Cancer Res. 83 (19), 3176–3183. https://doi.org/10.1158/0008-5472.CAN-22-3210. PMID: 37556505; PMCID: PMC10592143.