Investigating the Structural and Conformational Effects of G12D-K-RAS Oxidation through MD Simulations

دراسة الآثار الديناميكية والتكوينية الناتجة من أكسدة بروتين (RAS) من خلال المحاكاة الديناميكية للجزيئات

Shimaa Salamh

Supervisor: Prof. Abdallah Sayyed-Ahmad

August 2021

Investigating the Structural and Conformational Effects of G12D-K-RAS Oxidation through MD Simulations

دراسة الآثار الديناميكية والتكوينية الناتجة من أكسدة بروتين (RAS) من خلال المحاكاة الديناميكية للجزيئات

Shimaa Salamh

August 2021

Thesis committee:

Prof. Abdallah Sayyed-Ahmad (Principal advisor)

Prof. Wael Karain (Member)

Dr. Hazem Abu Sara (Member)

This thesis was submitted in partial fulfillment of the requirements for the Master's Degree in Physics from the Faculty of Graduate Studies at Birzeit University, Palestine.

Investigating the Structural and Conformational Effects of G12D-K-RAS Oxidation through MD Simulations

By

Shimaa Salamh

Accepted by the Faculty of Graduate Studies, Birzeit University, in partial fulfillment of the requirements of the degree of Master of Physics.

Thesis committee:

_____

Abdallah Sayyed-Ahmad Ph.D. (Principal advisor)

_____

Wael Karain Ph.D. (Member)

_____

Hazem Abu Sara Ph.D. (Member)

August, 2021

الإهداء

إلى منبع طموحي وملهميَ، إلى من وضعوني على طريق الحياة وكان دعاؤهم سر نجاحي
أمي وأبي

إلى رفيق الحياة ودفء المشاعر، إلى من رافقني في مشواري وكان لي سندا وعونا
إليك مصطفى

إلى سعادتي وسر بسمتي ، أبنائي
حازم ويامن

إلى أشقاء الروح، إلى من استمد منهم عزتي وإصراري
أخوتي وأخواتي

إلى أستاذتي ودكاترتي .. الى صديقاتي وزميلات العمل

إليكم جميعا أهدي بحثي هذا

Acknowledgment

Foremost, I'd like to express my thanks and appreciation to my advisor Prof. Abdallah Sayyed-Ahmad for his support, patience, immense knowledge, and invaluable guidance throughout this work. It was a great honor to work and study under his supervision. I would also like to thank all of the faculty members at the physics department for all of the knowledge and experience I gained during my studies. A special thanks to my thesis committee Prof. Wael Karin and Dr. Hazem Abusara for reviewing my thesis.

Abstract

Ras proteins are the main members of human Ras small GTPases that mediate a wide variety of cellular processes. They work as molecular switches in regulating many fundamental signaling pathways that are responsible for cell proliferation, differentiation, and survival. The hyperactivation of Ras signaling can occur directly through Ras mutations and is thought to be a key factor in cancer development. The mutations in Ras proteins are among the most powerful oncogenic drivers in 30% of all human cancers and are involved in tumor initiation and maintenance. A growing body of evidence suggests that Ras proteins could be regulated by redox reactions of cysteine residues found in the conserved redox-sensitive sequences known as the NKCD (Asn116-Lys117-Cys118-Asp119) motif of the G-domain. This redox signaling is a type of signal transduction is critical to physiological and pathological processes and occurs when Cysteine118 in Ras protein is oxidized in a reversible manner. In this study, we utilized all-atom Molecular Dynamics Simulations to investigate the structural and conformational effects of Cys118 oxidation on G12D-K-Ras. We have found that the oxidized variant is more dynamic than G12D-K-Ras, and the Cys118 oxidation alters the conformation of the nucleotide-binding site (the switches regions) of G12D-K-Ras, as well as perturb the conformational equilibrium between Ras active and inactive states.

ملخص

تعتبر بروتينات Ras من الأعضاء الرئيسية في مجموعة بروتينات GTPases البشرية الصغيرة، والتي تتوسط مجموعة واسعة من العمليات الحيوية في الخلية. إنها تعمل كمفاتيح جزيئية في تنظيم العديد من مسارات الإشارات الأساسية المسؤولة عن تكاثر الخلايا وتمايزها وبقائها على قيد الحياة. في بعض الأحيان يحدث تنشيط مفرط لإشارات Ras من خلال حدوث الطفرات، ويُعتقد أنه عامل رئيسي في تطور السرطان. تعد الطفرات في بروتينات Ras من أقوى العوامل المسببة للأورام في السرطانات البشرية. حيث أنه في ٣٠٪ من حالات السرطان التي تصيب الإنسان، يعتبر Ras هو المحرك الأساسي لبدء حدوث الورم وبقاءه. تشير العديد من الأدلة إلى أنه يمكن تنظيم بروتينات Ras عن طريق تفاعلات الأكسدة والاختزال للحمض الأميني Cysteine الموجود في المنطقة المحفوظة والمعروفة باسم NKCD. يعتبر نقل الإشارة بواسطة التأكسد والاختزال نوع من التأشير في الخلايا، ويحدث عندما يتأكسد Cys118 في بروتين Ras بطريقة قابلة للعكس، مما يؤدي إلى التعديل على بنية البروتين والذي قد يؤثر على استقرار البروتين، ونشاطه، وتمركزه، وكذلك التفاعل بين البروتينات، وقد وجد أنها ذات تأثير مباشر على العمليات الفسيولوجية والمرضية في الخلية. في هذا البحث استخدمنا المحاكاة الديناميكية الجزيئية لدراسة التأثيرات الهيكلية والتشكيلية الناتجة عن أكسدة الحمض الأميني Cys118، وكذلك لتحليل الحالة الديناميكية لكلا البروتين: G12D-K-Ras والبروتين المؤكسد . لقد توصلنا في هذه الدراسة إلى أن البروتين المؤكسد أكثر ديناميكية من بروتين G12D-K-Ras ، وأن أكسدة الحمض الأميني Cys118 تغير من البنية التكوينية و التشكيلية في موقع ربط النيوكليوتيدات (مناطق المفاتيح) لـبروتين G12D-K-Ras ، بالإضافة إلى اضطراب في البنية التشكيلية المستقرة له بشكل عام.

Table of Contents

List of Figures:

Chapter 1: Introduction

1.1. Ras Proteins

Ras proteins are the founding members of the large superfamily of small guanosine triphosphatases (GTPases) that comprise over 150 human proteins and divided into five major branches: Ras, Rho, Rab, Ran and Arf [1], [2]. Ras proteins have three isoforms H-Ras, K-Ras, N-Ras [3], which have an amino acid sequence consisting of 188-189 amino acids depending on the splice variant. The structure of Ras has two domains: a soluble catalytic domain (G-domain: amino acids 1-166), and a membrane anchoring hypervariable region (HVR: amino acids 167–189). The sequence of the G-domain of the different isoforms is highly conserved, while that of the hypervariable region is significantly different [4], [5], [6].

Ras isoforms share a common biochemical mechanism through which they act as molecular switches allowing them to regulate many fundamental signaling pathways responsible for cell proliferation and survival [7]. Biological activities of Ras are governed by a GDP/GTP cycle through which its affinity for downstream effectors is modified due to conformational changes depending on GDP/GTP binding [8]. The signal transduction is accomplished through reversible GTP binding, while the inactive form is bound to GDP [9], Ras's association with GTP or GDP is regulated by two enzymes: guanine nucleotide exchange factors (GEFs), which increase the rate of GDP dissociation, and GTPase-activating proteins (GAPs), which speed up the slow intrinsic rate of GTP hydrolysis [10]–[12], The structural differences between GDP-bound Ras (inactive state) and GTP-bound Ras (active state) are primarily found in highly dynamic regions known as Switch I (SI: residues 25–40) and Switch II (SII: residues 60–75), which are required for Ras interaction with both upstream and downstream partners [10], [12], [13].

1.2. G12D-K-Ras

GTPase K-Ras belongs to Ras superfamily, so it is a signal transducer protein that is involved in a variety of cellular signaling processes [14]. It is attached to the plasma membrane's inner leaflet, where it serves as regulatory switches, relaying signals from active receptors to cytoplasmic signaling cascades [15], [16]. The hyperactivation of Ras signaling, can occur directly through Ras mutations or indirectly through other proteins in Ras pathways, which is considered a key factor in cancer development [14]. The mutant K-Ras protein is one of the most common drivers of human cancer, accounting for 85% of all mutated Ras proteins found in human tumors [17]. When the K-Ras gene mutates, it gains oncogenic properties which impede GTP hydrolysis, resulting in the Ras molecules being permanently activated, so it appears to be causally involved in the development of a variety of human cancers [18], [19].

Natural Ras oncogene mutations have been identified in codons 12, 13, 59, and 61 [19], but the most common mutations detected in the K-Ras gene of cancer cells are at locations 12 and 13 [20]. Around 80% of patients have codon 12 mutations, while 18% have codon 13 mutations, and a much lower frequency (< 5%) at codons 59, 61 [21]. These allelic mutations cause amino acid substitutions, such as Gly to Asp, Ala, Arg, Ser, Val, or Cys in codon 12 [22], so it's written as G12X, with X is the new amino acid. These changes are close to the GTP binding site in the protein [18], which leads to conformational changes that make the protein stick in its active state for a significantly longer period of time than its nonmutated counterpart [23]. A cartoon representation of G12D K-Ras catalytic domain is shown in Figure (1.1).

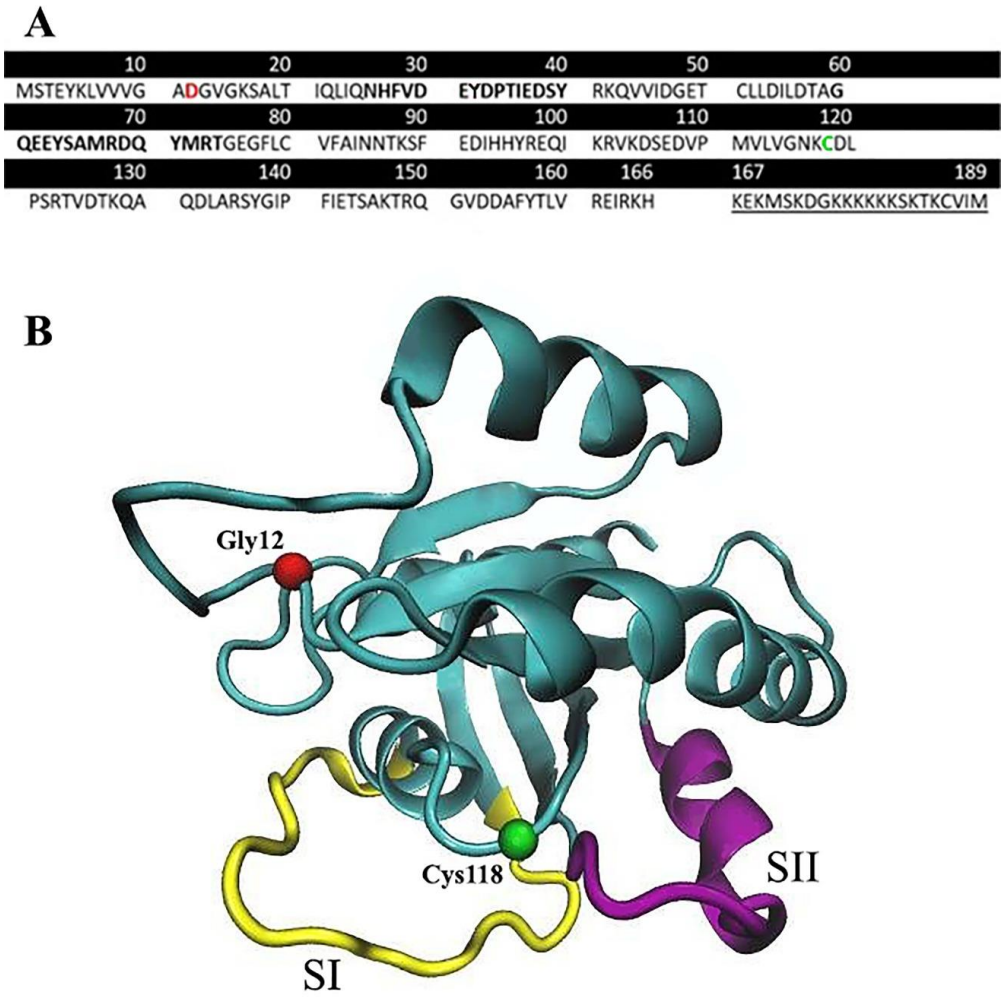

Figure 1.1 (A) G12D-K-Ras amino acid sequence of the catalytic domain (residues 1-166) and HVR region (residues 167-189), the residues 12, 118 are in red, and green respectively. (B) Cartoon representation of G12D-K-Ras catalytic domain (PDB 4DSO) with the location of mutations studied in this work, residue 12 colored in red, and green for residue 118. The Switches regions (SI and SII) are highlighted in yellow and purple respectively.

1.3. Oxidation of Ras Proteins

The small GTPase family contains a few redox-sensitive members (H-Ras, N-Ras, K-Ras, and some Rab proteins). Their NKCD (Asn116-Lys117-Cys118-Asp119) conserved redox-sensitive sequences, are found in the G-domain and are nearly identical in all proteins that contain this motif. The redox agents action on these redox-sensitive GTPases is similar to that of guanine nucleotide-exchange factors (GEF's) [4], [9]. In Ras proteins, the nucleotide-

binding regions include the two switches: SI which interacts with the nucleotide base, ribose, and γ-phosphate in the GTP-bound form. and SII which along with the P-loop (residues 10-17) interacts with the phosphate group of the bound nucleotide. On the other hand, the NKCD motif (residues 116-119) interacts with the bound nucleotide's guanine base. As a result, a mechanical disruption of the binding interactions between the nucleotide and these small motifs is needed to GDP allowing Ras activation [12], [13]

Redox signaling is a type of signal transduction that occurs when cysteines in proteins are oxidized in a reversible manner [4], which makes cysteine oxidation is a kind of PTM in Ras superfamily GTPases, that is frequently described as a new and emerging method of GTPase regulation [24]. Cys118 residue in NKCD motif of Ras proteins is found to be target site of the redox agents that mediate guanine nucleotide dissociation. Many studies have shown that Cys118 oxidation affects protein activity, stability, and localization, as well as protein-protein interactions. Meanwhile, cysteine residues at positions 80, 181, 184, and 186 may also contribute to the modulation of these actions. [4], [9], [13], [25].

1.4. Redox Agents:

Reactive oxygen species (ROS) and reactive nitrogen species (RNS) have been identified as cellular redox signaling agents [13]. ROS are a heterogeneous group of chemically reactive ions and molecules formed by the reduction of molecular oxygen O_2, including superoxide ($O_2^{\bullet-}$) and hydrogen peroxide (H_2O_2). While nitric oxide ($NO^{\bullet}$), and nitrogen dioxide ($NO_2^{\bullet}$), sometimes considered ROS because of the existence of a moiety of oxygen. ROS and RNS are produced in the cell as a result of various cellular processes. For example, H_2O_2 is produced as part of the electron transport chain in the mitochondria, and NOSs (Nitric Oxide Synthases) produce nitric oxide [26]–[28]. According to research, both $O_2^{\bullet-}$ and •NO can diffuse to neighboring cells across the cell membrane, allowing them to function as a signaling agent [26]. As a result, ROS and RNS have been found to act as second messengers [29].

1.5. Cysteine Redox Chemistry

Certain proteins are subjected to reversible chemical modifications in response to changes in localized redox potential. Due to the thiol groups (–SH) on cysteines, which are regarded a susceptible redox-sensitive targets, cysteine is widely utilized as a nucleophile in enzyme active sites, making it one of the most reactive members of nature's standard stable of amino acids [30]. The side chain of cysteine is easily oxidized to produce a variety of products due to the multiple oxidation states of the sulfur atom. However, not all cysteines contains thiols that are equally intrinsically reactive, since the thiol group (-SH) (the protonated form) is not particularly reactive, whereas the thiolate anion (-S) (the deprotonated form), is nucleophilic due to its abundance of available electrons [30], [31]. The cysteine thiol's specific reactivity in the protein is influenced by its microenvironment, with local polarity and interactions with neighboring residues influencing its pKa and redox potential, and since most thiols have a pKa of 8–9, indicating that they are nearly fully protonated at physiological pH and thus less vulnerable to oxidation [27]. However, Protein thiols with low pKa, especially those ionized at physiological pH, are commonly referred to as "reactive cysteines", so low pKa is an important factor in oxidization susceptibility [32]. For Cys118 side chain, the pKa value is likely to be lower than 8.3. This analysis suggests that, at least in part, at physiological pH (i.e. pH 7.4). the Cys118 side chain is most likely to be presented as a form of RS^-.

The thiol group goes through a number of oxidative post-translational modifications, which can result in sulfenic acids, sulfinic acids, sulfonic acids, S-nitrosothiols, sulfenamides, disulfides, and persulfides, as well as intramolecular disulfide bridges and intermolecular disulfides with small molecules like glutathione [27]. In general, sulfenic acids (R-SOH) are produced in proteins by H_2O_2 oxidation of cysteine's thiolate side chain. A sulfenic acid can be oxidized once more to produce a hyperoxidized sulfinic acid cysteine ($R\text{-}SO_2H$). As the reactive species levels increasing, cysteines can be oxidized further to form sulfonic acid (R-

SO_3H). While sulfenic acids can be reversed by the thioredoxin and glutathione enzyme systems, which act as an antioxidant defense system to protect against irreversible oxidation and can facilitate protein S-glutathionylation. In certain proteins, the sulfinic state can only be reversed enzymatically. Meanwhile the modification of sulfonic acid is thought to be irreversible and could be responsible for protein damage instead of signaling [27], [31], [33].

NO• modification of cysteine residues results in S-nitrosylation [34], which can occur through a variety of mechanisms dictated by the cellular environment, the most common and chemically readily available pathways involve the •NO-O_2 reaction products: (•NO_2), (N_2O_3). •NO_2 and N_2O_3 can both react with cysteine thiols to form S-nitrosothiols [35], (Figure 1.2).

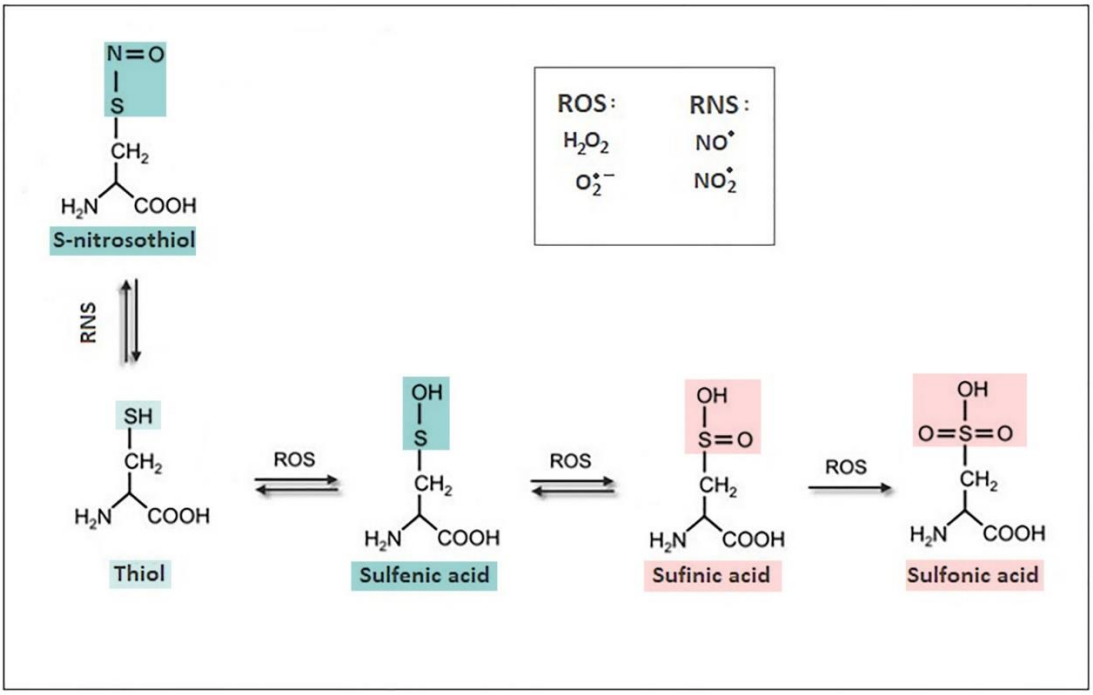

Figure 2.2: Possible oxidative post-translational modifications of cysteine. The thiol (-SH) group of cysteine residues can be modified by ROS or RNS to produce various oxidized derivatives. Sulfenic acid and S-nitrosothiol could be reduced by some reductants like glutathione and thioredoxin. While Sulfinic acid could be reduced specifically by sulfiredoxin in certain proteins. The modification of sulfonic acid is irreversible. Blue represents reversible processes, while Red represents irreversible processes.

Most of the oxidation reactions are irreversible, so the thiol modifications usually play only a limited role in regulating redox-sensitive proteins [36]. S-thiolation and S-nitrosation,

on the other hand, are common reversible post-translational modifications in proteins that play an important role in signal transduction, and may constitute a cell's protective/adaptive strategy [36], [37].

Many studies have demonstrated that NO can regulate Ras activity through the formation of the intermediate thiyl in Ras protein (Ras-S•) during the process of S-nitrosylation of Cys118's thiol group, which promotes the slow intrinsic dissociation of guanine nucleotide substrates from Ras. Lander and colleagues were the first to discover that NO can activate Ras in their investigation by increasing Ras GDP dissociation and stimulating pathways downstream of Ras [26], [38], [39]. Similarly during S-glutathionylation of Ras Cys118 residue, a Ras thiyl intermediate (Ras-S•) can also be formed which stimulating Ras activation through the slow intrinsic dissociation of GDP and GTP exchange [26], [40].

1.6. Physiological and Pathological Implications of oxidation of RAS proteins

Historically, reactive nitrogen species (RNSs) and reactive oxygen species (ROSs) were regarded as destructive oxidants capable of reacting with and damaging many biological macromolecules such as DNA, lipids, and proteins (8), but in 1995, the first evidence of RNS-induced RAS activation was found in human T cells [4]. Over the last few years, a growing number of studies have shown that the intracellular ROS production is heavily controlled, and that these redox agents also act as second messengers in healthy cells, where they take part in variety of signal transduction pathways and play an important role in physiological processes such as cell proliferation, differentiation, vasodilation, and migration [9], [27]. The redox agents as a downstream and/or upstream regulators of the redox-sensitive proteins play a significant role in cellular signal transduction. However, redox agents' dysregulation of small GTPases, or small GTPase misregulation of redox signaling, alters cellular signaling pathways. These changes frequently lead to different pathologies such as cancer and other diseases [9],

[13]. Also, the oxidative stress produced by a high level of ROS at the wrong place and wrong time causes cellular malfunction and apoptosis [34].

Many malignant cell types show an unusual redox metabolism, which includes antioxidant enzymes deregulation, and increased the production of reactive oxygen species (ROS) [36]. Because ROS signaling functions and toxic effects are concentration-dependent pathologies, tumor cells must actively control their ROS levels by increasing their own antioxidative capacity to avoid cell death. [26].

Ras is thought to be the most common oncogene in human cancer because activated Ras mutations are found in 30% of all human tumors [13]. As a result, cancer is one of the most common diseases caused by a redox agent's misregulation of Ras activity. Numerous studies have found that cancer is largely caused by Ras-redox signaling misregulation paired with a variation of Ras downstream cellular signal transduction cascades [9], [36]. Interestingly, It has been observed that the topical application of a NO-releasing agent (-+)-(E)-4-methyl-2-[(E)-hydroxyimino]-5-nitro-6-methoxy-3-hexenamide) to Sencar mice increases tumor-initiating activity by inducing a mutation in H-Ras at amino acid 13 and 61 [41]. Another research groups (Huang et al, 2014, 2015) try to investigate the role of the redox sensitive Cys118 in tumorigenesis in mice. They found that the loss of redox dependent reactions with Cys118 affects urethane-induced lung tumorigenesis as well as oncogenic H-Ras-driven tumorigenesis. Whereas in a mouse model of tumorigenesis, a Cys118 mutation (Cys to Ser) inhibited the growth of lung tumors [4], [42], [43]. On the other hand, some studies focused on understanding how H_2O_2 affects the tumor microenvironment. In these studies, breast cancer cells were co-cultured with cancer-associated fibroblasts, interestingly they found tumor H_2O_2 led to increasing ROS in cancer-associated fibroblasts [44]. These investigations are clearly critical in understanding the functional significance of cysteine oxidation in cancer.

Chapter 2: Methods

Proteins are flexible molecules that change shape and conformation as a result of interactions with other proteins or chemical modifications such as phosphorylation or oxidation (as in our case) [45], [46]. So the ability to track these changes is critical for understanding the structural conformational effects that result from these modifications. MD simulations, which are relies on the strict formalism of molecular physics, maybe the most accessible and appropriate method for modeling the protein motions at atomic level, and can trace and simulate the conformational changes in proteins [47].

In this study, we're interested in studying the structural and conformational effects of cysteine 118 oxidation by reactive oxygen species (ROS) that modifies the thiol group (-SH) into sulfenic acid (-SOH). This oxidative post-translational modification is known as cysteine sulfenylation, which is a reversible mechanism involved in Ras signaling as discussed in the first chapter.

To achieve this purpose, all-atom Molecular Dynamics (MD) Simulations of wild type thiol (C118-SH) G12D-K-Ras, and its oxidized mutant (C118-SOH) G12D-K-Ras were performed for 1 μs using NAMD2.11 [48] with CHARMM27 empirical force field and cMAP dihedral angle correction [49]. Then the trajectory files were visualized and analyzed using the trajectory analysis module in the Visual Molecular Dynamics (VMD) software [50], Bio3D package [51], Xm-Grace Visualization tool [52], and some inhouse Tcl scripts executed through the Tk Console of VMD.

2.1 Molecular Dynamics Simulations

Computer simulations are performed in the attempt of better understanding the features of molecular assemblies in terms of the microscopic interactions between them and their structures [53], one of the main simulation techniques is Molecular Dynamics Simulation which is regarded as a very effective tool for understanding biomolecular processes [54]. MD

simulations aim to predict how each atom in a protein or any other molecular structure will move over time using fundamental Newtonian physics approximations [55], [56], It bridges the gap between the macroscopic and the microscopic length and time scales, allowing it to be used to discover some bulk properties of a model system more easily than experiments on actual systems [53], [57]. To prepare a computer model of the molecular system, MD simulation is frequently used in conjunction with many experimental structural biology techniques, such as cryoelectron microscopy (cryo-EM), X-ray crystallography, electron paramagnetic resonance (EPR), forster resonance energy transfer (FRET), and nuclear magnetic resonance (NMR) [56], [57].

In MD simulations Newton's equations are solved numerically, which for a simple atomic system may be written:

$$m_i \ddot{\vec{r}}_i = -\vec{\nabla}_i U \qquad (2.1.)$$

So, we need to be able to calculate the forces acting on the atoms, which are derived from potential energy, $\vec{r_i}$ represent the complete set of 3N atomic coordinates, and m_i is the mass of each atom [53], [58].

The interatomic forces in MD simulation are approximated using a model known as a molecular mechanic's force field, which replaces the true potential with a simplified model designed to fit the results of quantum mechanical calculations and, typically, to certain experimental measurements on the one hand, and to be evaluated quickly on the other.

A common expression of the force field is like this:

$$U = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 \quad (2.2.)$$

$$+ \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\varphi - \delta)]$$

$$+ \sum_{improper} \frac{k_{imp}}{2} (\omega - \omega_0)^2$$

$$+ \sum_{LJ} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) + \sum_{elec} \frac{q_i q_j}{r_{ij}}$$

The first four terms are caused by interactions between chemically bonded atoms: bond stretching, angle bending, the torsional dihedral angles, and sometimes the improper torsions; where r_0 and θ_0 are the equilibrium bond length and angle respectively, φ : the torsional angle, δ is the phase, n denotes the number of minima or maxima between 0 and 2π or the multiplicity, and V_n determines the potential barrier's height. For the improper torsion, ω is the improper angle which referring to the deviation from planarity. k_b, k_a, k_{imp} are the bond, angle, and improper dihedral constants.

The last two terms describe the non-bonded forces that arise as a result of Van der Waals and repulsive interactions, as defined by the Lennard-Jones (6-12) potential, and electrostatic interactions, where **σ** is the diameter, **ε** is the well depth, and r_{ij} the inter-particle distance [56], [58], [59].

Examples of force fields that are commonly used in MD simulations to study proteins are CHARMM [60], AMBER [61], OPLS [62], GROMOS [63].

2.2 Protein Structure Preparation

The protein models for MD simulations were created using high-resolution crystal structures of G12D-K-Ras (PDB ID: 4DSO). The model containing sulfenic acid at Cys118 was

constructed by mutating Cys-SH to Cys-SOH using CHARMM General Force Field (CGenFF) [64] in CHARMM-GUI [65], where the force field parameters for the SOH form (cysteine sulfenic acid) were previously published in Heppner et al work [66]. The parameterization file has been added to the appendices.

In both cases, the guanosine triphosphate GTP was used to replace the bound guanosine-diphosphate-monothiophosphate (GSP), and all other co-crystals were removed except for waters and Mg^{2+}. The protonation states of each amino acid residue were predicted assuming neutral pH by using the PROPKA program [67]. For Cys-SOH it could be deprotonated (Cys-SO^{-}) because it is a weak acid. In the context of small molecules, various Cys-SOH pKa estimates range from 6 to 10. Recent estimates for Cys-SOH pKa values in dipeptides and selected proteins range from 5.9–7.2 [66], implying that under physiological conditions, a significant fraction of Cys-SOH exists in its protonated state. According to this, we used the protonated form of Cys-SOH in our MD simulations. Then the resulting structure was solvated in a TIP3P water box with a buffering distance of 10 Å. Furthermore, the solvation system was supplemented with (Na^{+}) and (Cl^{-}) ions for neutralizing and preserving a physiological concentration (0.15 M). An illustration of the final system configuration is shown in Figure (2.1).

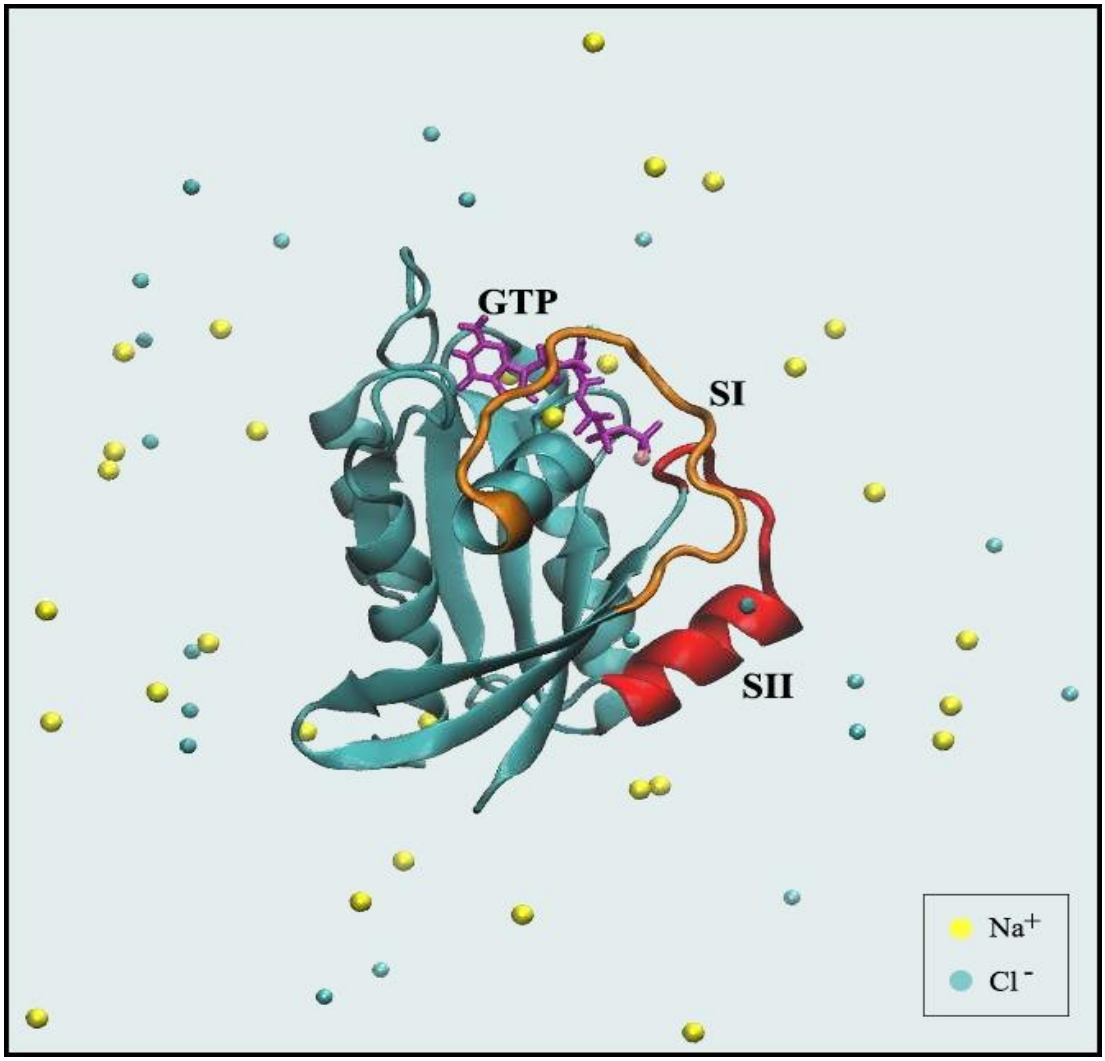

Figure 2.1*:* A snapshot from the MD simulation showing the catalytic domain of the protein colored in cyan. The switches regions SI (residues 25-40) is in orange and SII (residues 60-75) is in red. The magnesium, sodium, and chloride ions are shown as pink, yellow, and blue spheres. With the bound GTP highlighted in purple.

2.3 MD Simulation Setup

The systems were minimized using 5000 steps of conjugate gradient approach, then were slowly heated up from 0 K to 310 K at constant volume while the protein and GTP heavy atoms were constrained by a harmonic force constant k = 4 kcal/mol·Å^2, which was gradually removed at constant pressure while the temperature was maintained at 310 K, prior to the isothermal-isobaric (NPT) ensemble production run with periodic boundary conditions. The

simulations were run with a time step of 2 fs, and covalent bonds containing hydrogen atoms were constrained using the SHAKE algorithm [68].

To explain the long-range electrostatic interactions, the Particle Mesh Ewald (PME) method [69] was used, with a grid density of about 1/Å, and the non-bonded interactions were gradually switched off between 10 Å and 12 Å, and cutoff at 14 Å. NPT simulation was carried out at physiologic value of T = 310 K and 1 atm, to be consistent with the in vitro experiments. Langevin dynamics [70] with a damping coefficient of 10 ps^{-1} were used to regulate the temperature and pressure. And to maintain constant pressure, the Nose-Hoover Langevin piston method [71] was used, with a piston period of 200 fs and decay time interval of 100 fs.

2.4 Techniques to Analyze MD Trajectories

2.4.1. Root-mean-square deviation (RMSD)

RMSD is a popular metric for calculating the average distance between the expected and original positions of all atoms in two protein structures [72] thus, determining the extent of difference in their three-dimensional coordinates. As a result; the smaller the RMSD, the more similar the two structures are [73]. Also, it can be used to determine a structure's or model's conformational stability during that simulation [74]. In our study, RMSD analysis is used to evaluate the conformational differences in the G12D-K-Ras and its oxidized isoform catalytic domain, over the simulation's time course with respect to the initial X-Ray structure calculated after removing translation and rotation of the protein backbone excluding the flexible switch regions.

RMSD values are presented in Å and calculated as follows [75]:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|\vec{r}_i(t) - \vec{r}_i^{ref}\right|^2} \qquad (2.3.)$$

Where N is the number of atoms in the protein structure, $\vec{r}_i(t)$ is the position of the i-th atom at a given time t, and $\vec{r}_i^{ref}$ refers to the corresponding reference position of the i-th atom. RMSD can be calculated for any type and subset of atoms, such as Cα atoms of the entire protein or Cα atoms of all residues in a specific subset for example binding pocket, or a loop.

2.4.2. Root mean square fluctuations (RMSF)

While RMSD measures average position differences over all atoms, RMSF is used to calculate the deviation in the position of each atom averaging over time [76]. Thus, the RMSD can be used to identify the fluctuations of different protein conformations across the entire MD trajectory, whereas the RMSF can be used to calculate the dynamic fluctuations of each residue around a reference conformation, which is usually the average position in the aligned structures. RMSF is typically linked to the flexibility of protein structures because it reflects each residue's mobility during the MD trajectory [77].

The RMSF is given as:

$$RMSF = \sqrt{\frac{1}{T}\sum_{t_j=1}^{T}\left|\vec{r}_i(t_j) - \vec{r}_i^{ref}\right|^2} \qquad (2.4.)$$

Where **T** is the trajectory time, $\vec{r}_i(t_j)$ is the atom **i**'s position at time $\mathbf{t_j}$, and $\vec{r}_i^{ref}$ is the reference position of the i-th atom.

2.4.3. Order Parameter of Backbone Amide Bonds (S^2)

Countless biological processes, including folding and assembly, catalysis and ligand binding, rely on information transfer via proteins conformational changes, which results in a net entropy change [78]. In many previous studies, a significant change was observed in the flexibility of the protein backbone between biologically relevant conformations, which indicates that the

ps-ns protein dynamics can contribute significantly to biological function through changes in conformational entropy [79], where the protein conformational entropy appears to be a controlling factor in some binding reactions. However, in determining protein conformational entropy changes, integrating NMR data with MD simulations has found to be particularly beneficial [80].

To explore the dynamics of the entire protein including the long time scale as well as the small fluctuations, we calculated each residue's NMR order parameter, S^2 as ensemble averages according to [81]:

$$S^2 = \frac{3}{2}\left[\langle\mu_1^2\rangle^2 + \langle\mu_2^2\rangle^2 + \langle\mu_3^2\rangle^2 + 2\langle\mu_1\mu_2\rangle^2 + 2\langle\mu_1\mu_3\rangle^2 + 2\langle\mu_2\mu_3\rangle^2\right] - \frac{1}{2} \quad (2.5.)$$

where $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, and $\boldsymbol{\mu_3}$ represent the $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ components of $\boldsymbol{\mu}$, respectively. S^2 is calculated as the average of all snapshots in the trajectory [5]

S^2 is considered as a good predictor of the protein backbone motions in computationally feasible timescales [82], which shows the equilibrium distribution of the vector μ(t) orientations in a molecular reference frame [83].

The magnitude of S^2 can range between 0 and 1, with lower S^2 values generally corresponding to larger amplitude internal re-orientational motions, whereas high values indicate that the bond vector motions are more restricted relative to the molecular frame [84].

2.4.4. The Internal and Total Dynamics Correlation Function

For getting additional information about flexible regions in proteins, and their in-cell conformations. the correlation function $C(t)$ is used to characterize the molecular motion. It provides the memory of how long a molecular conformation remains unchanged according to the relaxation of the N-H dipoles of their backbone [5], [84], [85].

The correlation function describing the dynamics of an N-H bond at different times defined as $\mathbf{C(t)} = \langle \boldsymbol{P_2}\,(\hat{\boldsymbol{\mu}}_t \,.\, \hat{\boldsymbol{\mu}}_0)\rangle$ where $\mathbf{P_2} = \frac{3X^2}{2} - \frac{1}{2}$ is the second Legendre polynomial, and $\hat{\boldsymbol{\mu}}_t$ is a unit vector along the N-H bond at time t, while the angular brackets indicate averaging over time [5], [86]. C(t) general shape is usually a rapid initial decay to a plateau value due to internal motions that occur on a picosecond time scale, followed by a much slower decay due to the overall tumbling motion on a nanosecond time scale [84].

When the overall protein motion is not taken into account, the function is called internal correlation function $C(t)_{int}$. [82], which is obtained by superimposing the trajectory frames onto a reference structure (the initial structure) excluding the switches [87].

2.4.5. Principle Components Analysis (PCA)

The principal component analysis (PCA) method is a widely used statistical approach for analyzing protein motions, and it has proven to be an effective tool for investigating conformational changes by describing the concerted atomic displacements [83]. The purpose of PCA in this context is to reduce the dimensionality of a multivariate dataset in a way that can determine the most significant dynamics of the system [88], [89]. PC analysis, which can be carried out with cartesian coordinates or dihedral angles, has proven to be an effective tool for highlighting significant conformational changes between structures [90]

PCA basically is a linear transformation that diagonalizes the 3N×3N covariance matrix which is constructed from atomic coordinates (Cartesian) after the translational and rotational motions in the MD trajectory have been removed, thereby removing the instantaneous linear correlations between coordinates. Mathematically,

$$\mathbf{C} = \mathrm{V}\boldsymbol{\Lambda}\mathrm{V}^{\mathbf{T}} \qquad (2.6.)$$

Where **C** is the data covariance matrix, **Λ** represents a diagonal matrix containing the eigenvalues, and **V** is the matrix containing the corresponding eigenvectors (PCs) which are ordered by decreasing corresponding eigenvalue [91], [92].

This diagonalization process produces a complete set of orthogonal modes (eigenvectors), the eigenvectors of such a matrix are the "best fitted" directions through points in configurational space generated by an MD trajectory. Each of them has a corresponding eigenvalue "variance" that describes a portion of the motion, with larger eigenvalues "variance" characterizing motions on larger spatial scales, and the first eigenvector having the highest eigenvalue (average square displacement) possible. As a result, it's been demonstrated that the first few principal components can accurately define a significant portion of the system's fluctuations. To visualize the results of the PCA analysis, the original data are projected onto a two-dimensional plane using a transformation matrix defined by two eigenvectors of interest (usually the first two principal components, PC1, and PC2). The projection of the trajectory onto a specific eigenvector emphasize the time-dependent motions performed by the components in the particular vibrational mode [74], [89], [90], [93], so the dynamics of the protein in this low dimensional subspace spanned by the first few principal components were termed "essential dynamics" [90].

In this study, we used the Bio3D package available in R programming language to perform the PCA analysis [51].

Chapter 3: Results and Discussions

In this chapter, the findings of the analyses concerning the structural and conformational changes in the catalytic domain of G12D-K-Ras and its oxidized counterpart both in GTP nucleotide bound states will be discussed in the context of the potential functional significance of sulfenylation modification on Ras activity. Therefore, a number of analyses were executed on the resulting MD trajectories. Firstly, the results of specific differences in the conformational fluctuations of the catalytic domains in general and the switches regions of both mutants are presented, then we investigate the switch regions dynamics and the overall dynamics and flexibility, also we monitored Cys118 residue conformations and sodium ion interaction with GTP by visual inspection of the MD trajectories.

3.1. Dynamics, Flexibility & Structural Stability Analysis of the Models

In order to understand the effects of cystein118 oxidation on the conformational dynamics and equilibrium structures of G12D-K-Ras. RMSD of the backbone atoms for both mutant and native structure (for the catalytic domain and the switches regions) were monitored during the simulation time. From Figure (3.1), it is observed that the RMSD for SOH- protein fluctuates initially, and then becomes stable from 400 ns onward, meanwhile, the value for G12D-K-Ras reaches the plateau quickly and remain stable throughout the simulation, suggesting the convergence of simulations. The average RMSD values of SOH-G12D and G12D-K-Ras are found to be 1.5 Å and 1.1 Å, respectively.

Since the oxidized counterpart was constructed based on G12D-K-Ras structure, perturbations and conformation rearrangements to the structure may be introduced by the mutations, reflected by the increased RMSD values. So, the secondary structure of the oxidized protein experienced larger conformational changes [94]. According to one potential model

mechanism for explaining the action of redox agents on redox-sensitive small GTPases. The end product of the redox agent's reaction with the redox-sensitive residue of small GTPases alters the conformation of the nucleotide-binding site of small GTPases. Consequently, the bound nucleotide is released from small GTPases. Another possibility is that a chemical reaction of a redox agent with the redox-sensitive GTPase residue disrupts the interactions of Ras nucleotide-binding. This perturbation, according to this explanation, leading to the dissociation of the bound nucleotide from small GTPases [13].

However, significant differences were observed in the dynamical behavior of the two switches regions, especially in SII. For SI the average RMSD of G12D-K-Ras ranges between 0.9 – 2.8 Å and its oxidized variant 1.2 – 3.7 Å, while in SII for G12D-K-Ras is 1.1 - 2.8 Å and 0.1 – 4.9 Å for SOH-G12D protein. The RMSD was observed to increase as a function of time for SII of C118-SOH when compared with G12D-K-Ras. The significant changes of RMSD suggesting global conformational modifications from the closed-form to the open-form.

The GTP-bound form can exist in a conformational equilibrium between states II and I. State II represents the active form, capable of executing downstream signaling via directly interacting with its effectors, whereas state I's affinity for effectors is 20 times lower than that of state II [8]. Ras GTP-bound state II, structurally corresponds to a closed-form conformation in which the two functional loops in the binding region SI and SII interacts with the GTP's γ-phosphate. While in the GTP-bound state I and GDP-bound forms, Ras adopts an open-form conformation characterized by a separation of SI from the guanosine nucleotide, resulting in enhanced flexibility of both switches. The instability of SII, which is caused by Gly60 dissociation from the guanosine nucleotide, is a common structural feature of state I that distinguishes it from state II [8], [95].

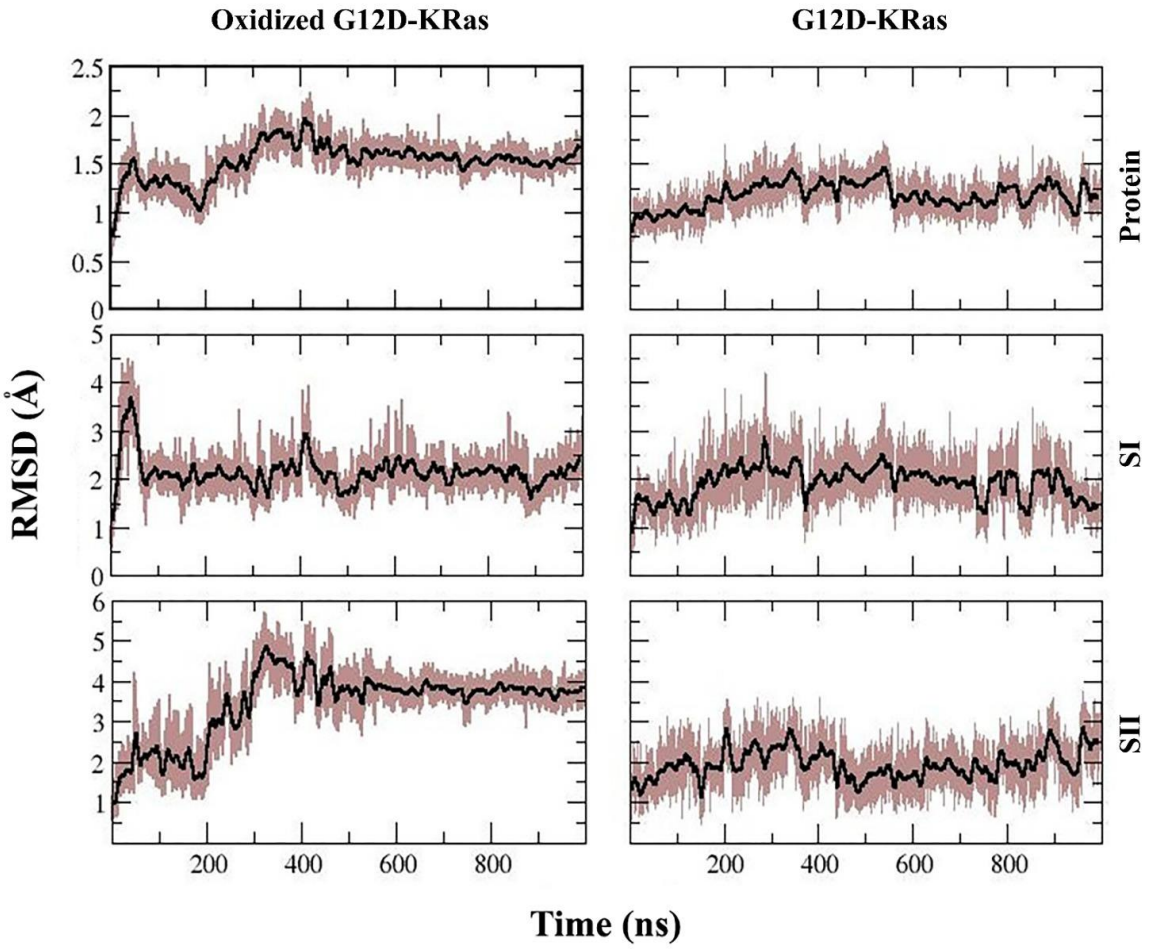

Figure 3.1: Time evolution of backbone, SI, and SII RMSD with respect to the initial structure of G12D-K-Ras and its oxidized variant, the RMSDs were evaluated after alignment excluding the flexible switch regions. The data were sampled every 100 ps (brown) with a running average every 10 ns (black).

To provide additional information pertaining to flexible regions, Figure (3.2) shows the RMSF of each residue in both systems. Where the RMSF represents the backbone atoms' dynamics, as higher values are associated with increased flexibility and mobility through the MD simulations [96].

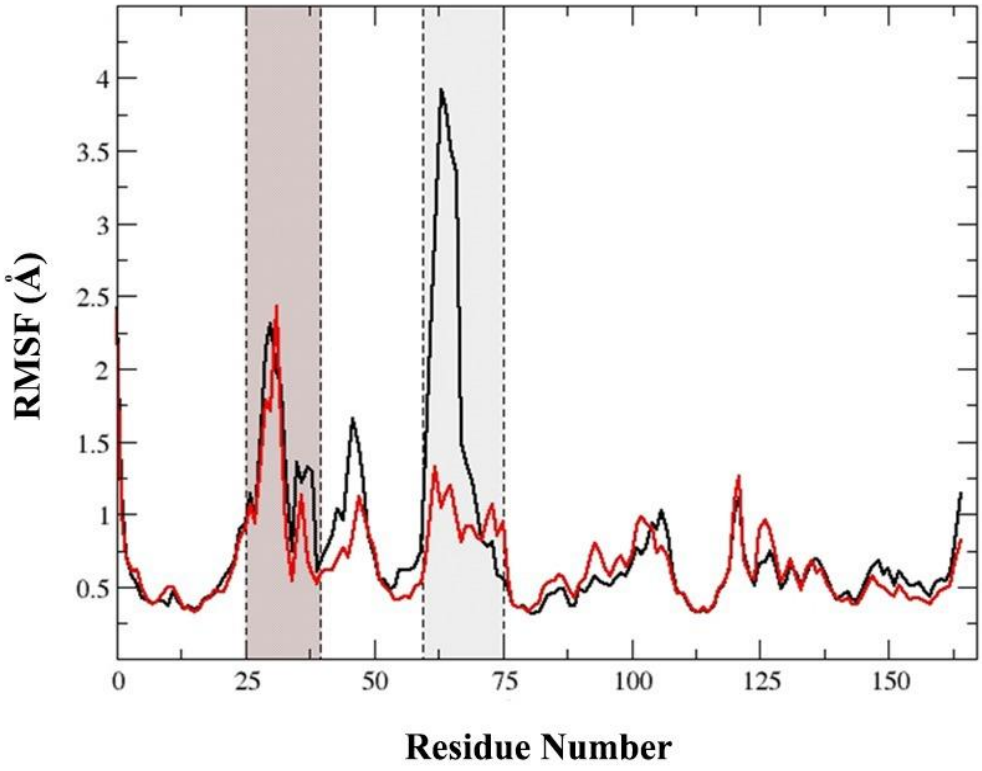

Figure 3.2: C_α -RMSF of G12D-K-Ras (red) and its oxidized variant (black), the RMSF is calculated after alignment excet for the flexible switch regions, which are highlighted in purple for SI and cyan for SII.

The effects of C118 oxidation of G12D-K-Ras appear clearly in the flexibility of SI and SII and β2-L3 regions, where the highest peaks on the graph are the mostly fluctuated during the simulation, Further, it is revealed in Figure (3.2) that both SI, SII, and β2- L3 (residue 41 - 64) regions show higher flexibility in the oxidized variant than the original system, with average RMSF of SI residues is 1.34 Å versus 1.12 Å in the unoxidized system, and the average RMSF of SII residues is 1.86 Å versus 0.97 Å. The structural disruption at one protein site affects the dynamics, structure, and biochemical properties at other sites that are of particular interest [94]. And since the nucleotide's nucleobase is coordinated by SI, L8 (which is contains Cys118), and L10. So, the oxidation of Cys118 makes minor modifications in that region that

may affect the crosstalk between SI and the allosteric lobe. Hence influencing the nucleotide exchange rate and intrinsic GTPase activity and effector binding [85].

The SI-β2 region has the most direct and strong contact with the effectors [8], therefore, the dynamics of this region are clearly critical for effector binding. The flexibility of SI caused by the residue Thr35 dissociation from the guanosine nucleotide which is detected in Ras-GTP state I, is associated with low affinity of effector binding [95], while the increased local flexibility reported in SII region may indicate that SII is involved in the binding and activation of these effectors by forming contacts with them in regions other than the canonical Ras binding domain [97].

Inter-lobe dynamics are generally determined by the bound nucleotide, with the inactive GDP-Ras form being more flexible and dynamic than the active GTP-Ras form. This implies that increased flexibility could promote GEF binding and (or) transmission between active and inactive states, or conversely [98]. In previous research a series of mutant Ras isoforms were studied using x-ray crystal structure analyses. This study has shown that the highly flexible nature of the switches regions, particularly SII, is responsible for the GDP/GTP and state I/state II transformations, which are facilitated by γ-phosphate positional changes [95]. This implies from these observations that the oxidation of G12D-K-Ras is might be responsible for a conformational transition in the protein.

3.2. Principal Component Analysis (PCA) of Catalytic Domain Trajectories

Generally, the specific functions of proteins are carried out through their collective atomic motions. Hence, it is used as a parameter to understand the stability of proteins and to quantitatively characterize the local fluctuations in the conformations. To probe how the mutation affect the dynamics of the backbone atoms, PCA is used to investigate the global motions of protein into a few principal components, characterized by eigenvalues and eigenvectors. A few low-frequency eigenvectors with large eigenvalues frequently account for

a significant portion of the overall macromolecules' fluctuations, if the motions are analogs, then the eigenvectors and eigenvalues from the individual trajectories should also be similar [74].

The first three PC subspaces describe 28.7%, 10.5%, and 7.8% of the total conformational displacement, Figure (3.3) shows the conformational sampling of G12D and the oxidized variant SOH-G12D in the essential subspace (PC1 and PC2). We only compared the first two PCs because they contribute the most to overall motion, in the two models, the first two eigenvectors account for ~ 40% of total motion, strongly dominating the overall variance, we will go over the conformational changes in the states by using these two PCs below. The fluctuations recorded by the first two principal components (PC_1 and PC_2) show that SOH-G12D samples a wide range of configurational space compared with G12D, However, the phase space sampled by the mutants is somewhat different. This disparity could be attributed to the differential dynamics at SII.

Applying PC analysis on the two systems by projecting the trajectory snapshots onto the plane formed by the first two principal components, reveals that the vast bulk of the motions for SOH-G12D are explored along PC1 between -10 to 10 Å while PC2 is populated between -10 to 5 Å, which showing unusual pattern on the phase space, a semicircle, or U-shape, relationship. While this result does not show any dominant large scale conformational changes within the system, it does show the more readily available degrees of freedom for thermal motion along the time scale under consideration [90]. Meanwhile, the unoxidized counterpart showed equivalent spread along the two PCs. This suggests that a significant alternation in the protein conformation is happened due to the oxidation.

The results clearly show that the G12D-K-Ras protein occupied a smaller region of phase space than the oxidized variant. Therefore, the PCA results indicate that the G12D-K-Ras

protein is more stable than the oxidized proteins, and Cys118 sulfenylation significantly altered structural stability and flexibility. The distribution also indicates the extent of fluctuation in each mutant: the narrow distribution of the G12D-K-Ras mutant represents its small fluctuations, whereas the wide distribution of the oxidized mutant represents its large fluctuations.

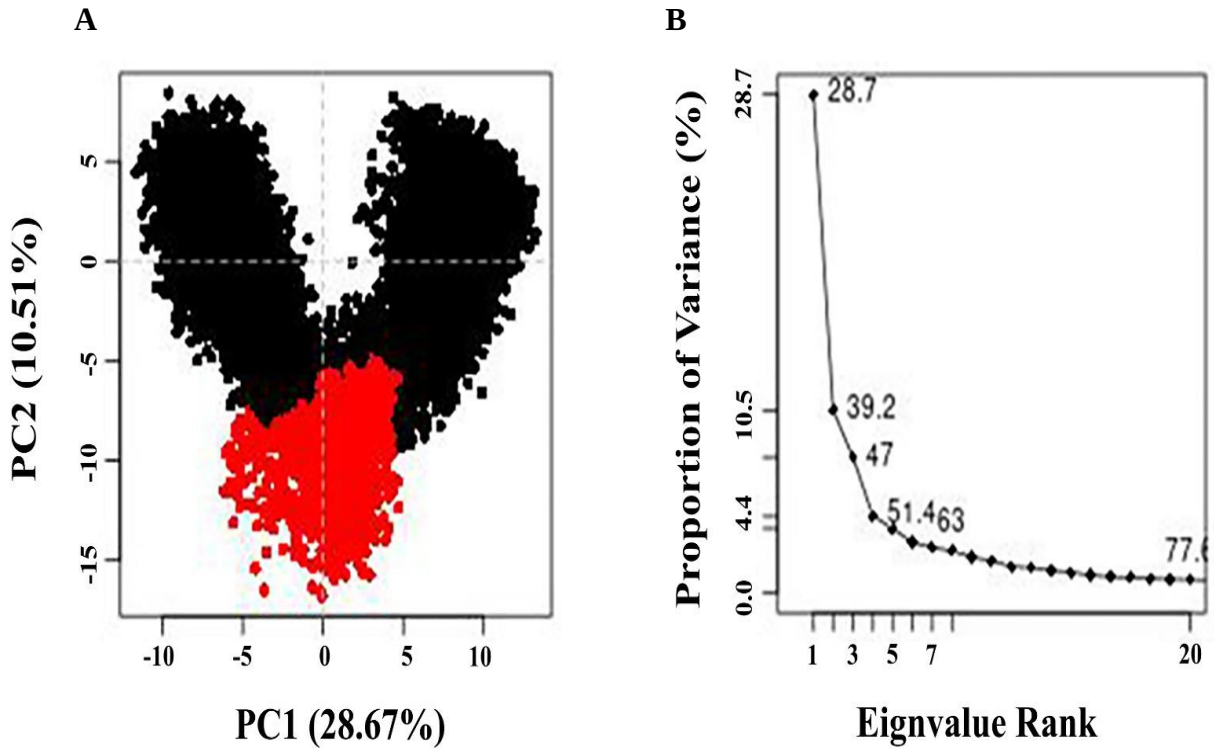

Figure 3.3: PC Analysis. (A) Global conformational dynamics of mutants G12D-K-Ras and its oxidized counterpart. The simulation trajectories are projected onto the space defined by the first two principal components (PC1 and PC2), the G12D-K-Ras is in red and the oxidized counterpart in black. (B) The eigenvalues are plotted vs. the eigenvectors indices of the covariance matrix.

Another strategy for interpreting PC analysis results is to assign individual residue contributions to the first PCs, these contributions are shown in Figure (3.4) for the first PC. Hence, closure looks to the mobility of the catalytic domain residues in terms of PC1, confirmed that the conformational changes were mostly induced in SII region of the oxidized counterpart, these results were correlated with the RMSF in the increase of the movement of

SII region. While in the unoxidized protein it can be seen that many motifs of the protein contribute, including the switches regions. While the chain endings contribute to both mutants. The contributions from the protein termini could be genuine because they do not share in a stabilized secondary structure [90]. In short, the PCA results are in agreement with the RMSD and RMSF findings, enhancing the validity of the performed analysis.

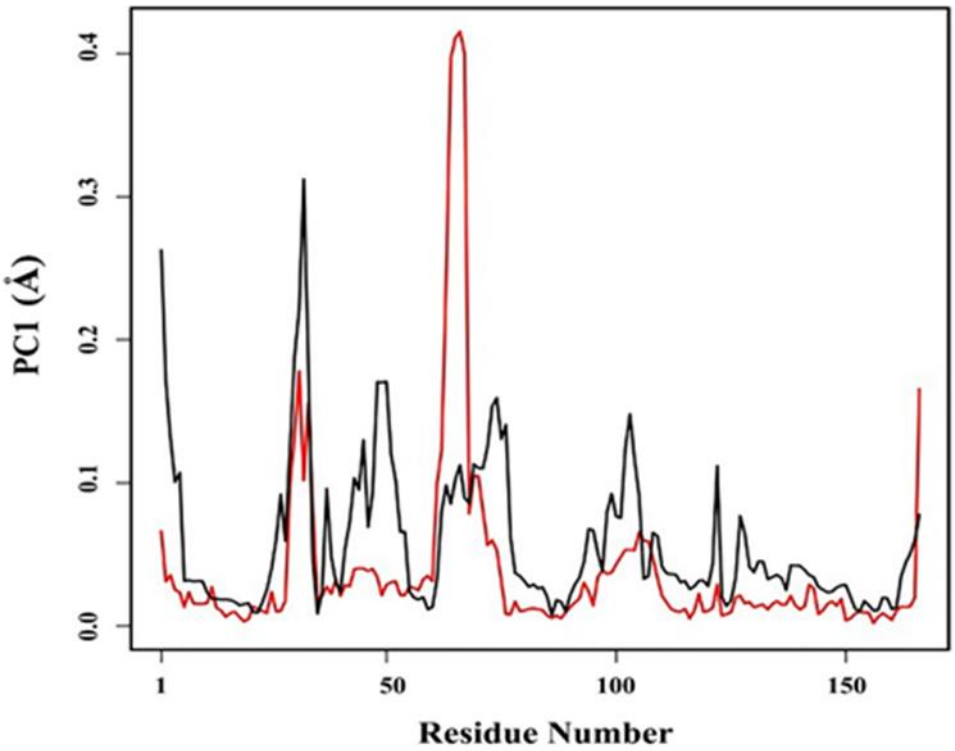

Figure 3.4: PCA loading or contribution of each residue of both mutants G12D-K-Ras (black) and the oxidized variant (red) to the first principal component (PC1).

3.3 Analysis of Cysteine 118 Side Chain Dihedral

Another feature which showed clear discrimination in the conformation of G12D-K-Ras and its oxidized variant is the dihedral angle of Cys118 residue χ (S-C_α-C_β-C_γ). For both trajectories, a visual inspection of the Cys118 side chain was performed. In the two mutant trajectories, the orientation of Cys118 (along with a neighboring Phe28 residue) differed.

Dihedral angle analysis for Cys118 residue was performed to quantify this change, as shown in Figure (3.5), the Probability Distribution of the dihedrals P(χ_1) reveals a large change in the orientation of the side chain of Cys118. For G12D variant the highly probable side-chain conformation can be expected at χ_1= - 60 °, On the contrary, the peak for SOH-G12D is nearly 170 °, indicating a high affinity for this conformation. And hence, the comparison of the Cys118 dihedral in G12D and its oxidized counterpart trajectories suggests the effects of Cys118 oxidation on the side-chain conformation.

According to the spatial arrangement of the redox-sensitive Cys118 side chain, and the Phe28 side chain. The Phe28 side chain faces the sulfur atom of the Cys118 side chain, and there is no residue between them. The distance between the Phe28 phenyl side chain's center and the Cys118 sulfur atom is 12 Å, which minimizes hydrogen, and ionic-bonding, and hydrophobic interactions, and not limiting the electron transfer between them [13], so the mutation on Cys118 side chain by sulfenylation may affect this interaction and change the spatial configuration of these residues resulting in a significant shift in the orientation of the side chain of Cys118.

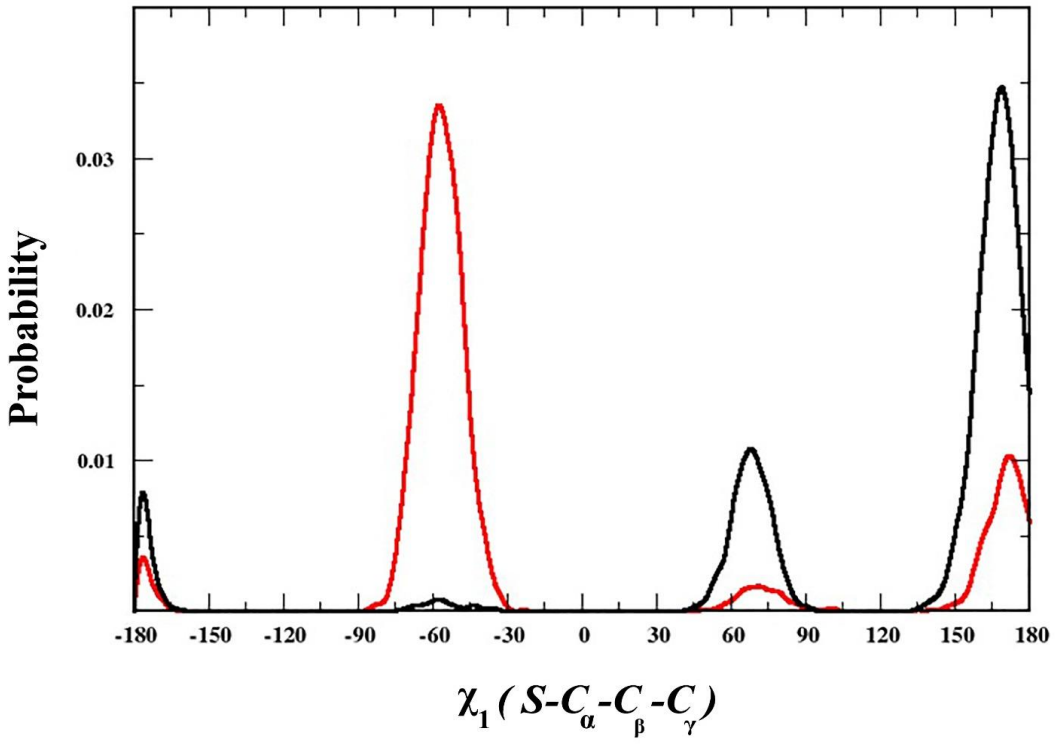

Figure 3.5: The dihedral angle χ_1 (S-Cα-Cβ-Cγ) probabilityof Cysteine 118 for G12D-K-Ras (black) and its oxidized counterpart (red).

3.4 Identification and Characterization of Ras Conformational States:

Based on the results of NMR studies, Ras can adopt two main conformational states when it's in complex with GTP according to SI conformation, a non-effector binding state I (inactive state) reflects to an open-conformation, and an effector binding state II (active state) which reflects to a closed-conformation that is also could found when RAS interacts with an effector protein. Ras state I have been proposed as in nucleotide exchange process is an intermediate state [99]–[101].

From previous studies it's found that in the P-loop mutations, such as G12D, shift the equilibrium between the two states toward the inactive state I [100], [102], that may impair

GTP hydrolysis, causing the G12D-K-Ras protein to adopt a permanent form and remain in an active GTP-bound state for a longer time, which explain the aggressive tumor phenotype caused by G12D-K-Ras mutant [21].

By monitoring the pocket distances of the binding regions of both mutants through utilizing the distance between the mass center of residues (12–13) and residues (32– 34) as a metric to distinguish conformations of state I and state II. we found that the GTP-binding pocket in the G12D-K-Ras protein was slightly more open than that of the oxidized variant, which adopts a more open SI conformation, indicates that the binding of GTP with the G12D-K-Ras mutant is less favorable when compared to that of GTP with oxidized K-RAS [21]. Figure (3.6).

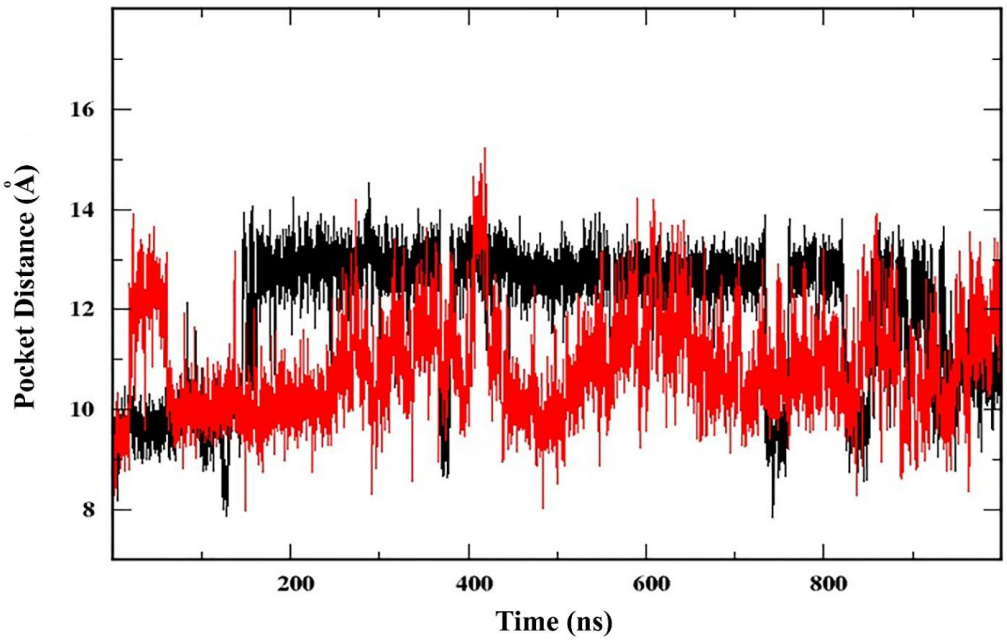

Figure 3.6*:* The pocket distances between the mass centers of residues 12-13 and 32-34 for G12D-K-Ras (black) and the oxidized variant (red).

Understanding the effects of Cys118 sulfenylation at protein GTP interaction and the conformational state is important since these conformational equilibria have a direct influence on Ras's interaction with its effectors. many researchers have explored state I's affinities to effectors in an effort to comprehend the properties of state I. They found the affinity to the effectors is significantly lower in Ras variants in state I [99].

In our simulation, SOH-G12D show an intermediate conformation which shift from state I by decreasing the opening of the pocket, but still in state I, many MD simulations showed that members of the small GTPase family in their GTP-bound form, may share the conformational equilibrium between the two states [100]. In a study done to comprehend the differences in conformational dynamics of G12 missense mutants in G12 K-Ras proteins, it covered the conformations of the cryptic state I, except for the closed state II conformation. According to these simulations, state I should be defined as an "ensemble of conformations" rather than a "single conformation", as it is most commonly described in the literature [14]. These conformational changes in the protein may affect its interactions with various downstream signaling transducers, such as the GTPase-activating protein (GAPs), which lead to an increase in the activity of the mutants [21].

GDP and GTP molecules bind to Ras proteins with high affinity due to strong interactions between the guanine nucleotide bases N_1-H and C_2-NH_2 and the Asp119 carboxyl side chain. This is confirmed by the observation that mutation of these residues considerably reduces guanine nucleotide-Ras binding affinity [13].

3.5 Backbone Relaxation Time and Order Parameter Analysis

The generalized order parameter (S^2) is used to calculate the degree of spatial restriction of internal motion. Therefore, to investigate the dynamics of the entire protein. we calculated the amide backbone order parameter (S^2) per each residue as an average over the entire trajectory. From Figure (3.7), the average order parameter $\langle S^2 \rangle$ is found to be equal 0.85, 0.80 for G12D

and SOH-G12D respectively, indicates that most amide bond vectors have a high motional restriction in their orientation, which shows a well-ordered structure of both mutants. In contrast, some residues like T2, G77, Q150, and SI, SII, and some other loop regions show increased flexibility as revealed by low order parameters (0.16 – 0.6). Importantly, amino acids involved in intramolecular hydrogen bonds were observed to have higher $\langle S^2 \rangle$ values than the other amino acids[84]. These findings are supported by the MD simulations as the backbone fluctuations show a local maximum for the residues in these regions.

However, differences in S^2 between the mutants are mostly observed in flexible regions. A significant increase in the flexibility is observed in SII region (residues 61 - 68) in SOH-G12D variant with very low S^2 values (0.11 – 0.31). Meanwhile, SI is already flexible, a drop in S^2 value in the residues (37 – 40) indicates that large-amplitude motions occur in these residues compared with the unoxidized variant. Slightly more mobile residues are also found in β2-L3 region in SOH-G12D, which is consistent with the previous observations.

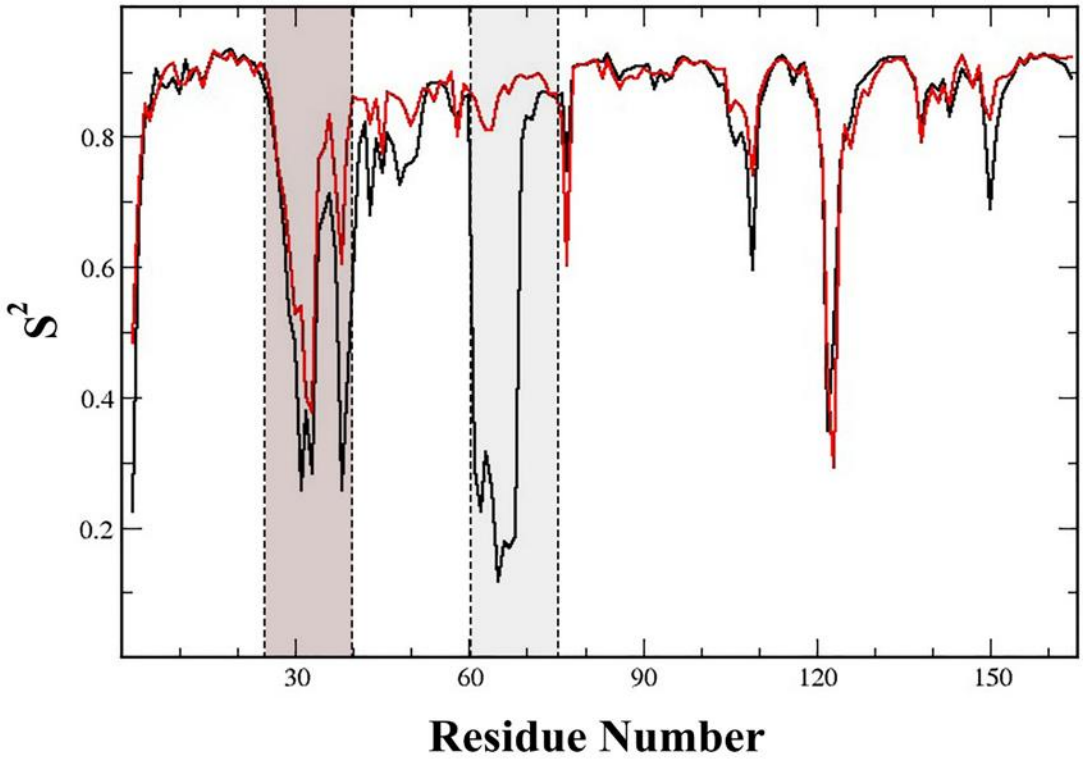

Figure 3.7: Backbone amide order parameter S^2, evaluated for the entire trajectory time of both mutants colored in red for G12DK-Ras and black for the oxidized variant.

To gain insight into the nature of switch regions dynamics, the total and internal correlation functions C(t), $C(t)_{int}$ respectively of the backbone amides were evaluated up to 1microsecond, but it could be meaningful for much shorter times [63]. As shown in Figure (3.7), the overall rotational motion C(t) of the switches regions decay very slowly over the trajectory time, toward zero within ~100 ns for both mutants of the two switches, and diverge from the curve of internal correlation function $C(t)_{int}$ beyond 10 ns, which represents the decorrelation of the bond vector direction caused by molecular tumbling, and the long-range dynamic processes cause the post-decay of $C(t)_{int}$. This behavior is a typical example of what would be predicted for residues undergoing rapid, small amplitude movements [5]. Now by looking at $C(t)_{int}$ curves, the internal correlation functions decay to a plateau value, that defines the square of the order parameter (S^2) [103]. As expected, the internal correlation function for

SII of the oxidized G12D-K-Ras rapidly decays to the smallest order parameter value, because this region samples a larger conformational ensemble, whereas the correlation functions of the SI and SII of G12D-K-Ras decay slower to the larger order parameter values.

Since both switches of SOH-G12D have faster relaxation dynamics within the time scale, suggesting that they are more flexible than the G12D-K-Ras variant. Meanwhile G12D-K-Ras exhibit slower relaxation SII than SI indicating that SI is more dynamic and flexible, the contrary has happened for SOH-G12D.

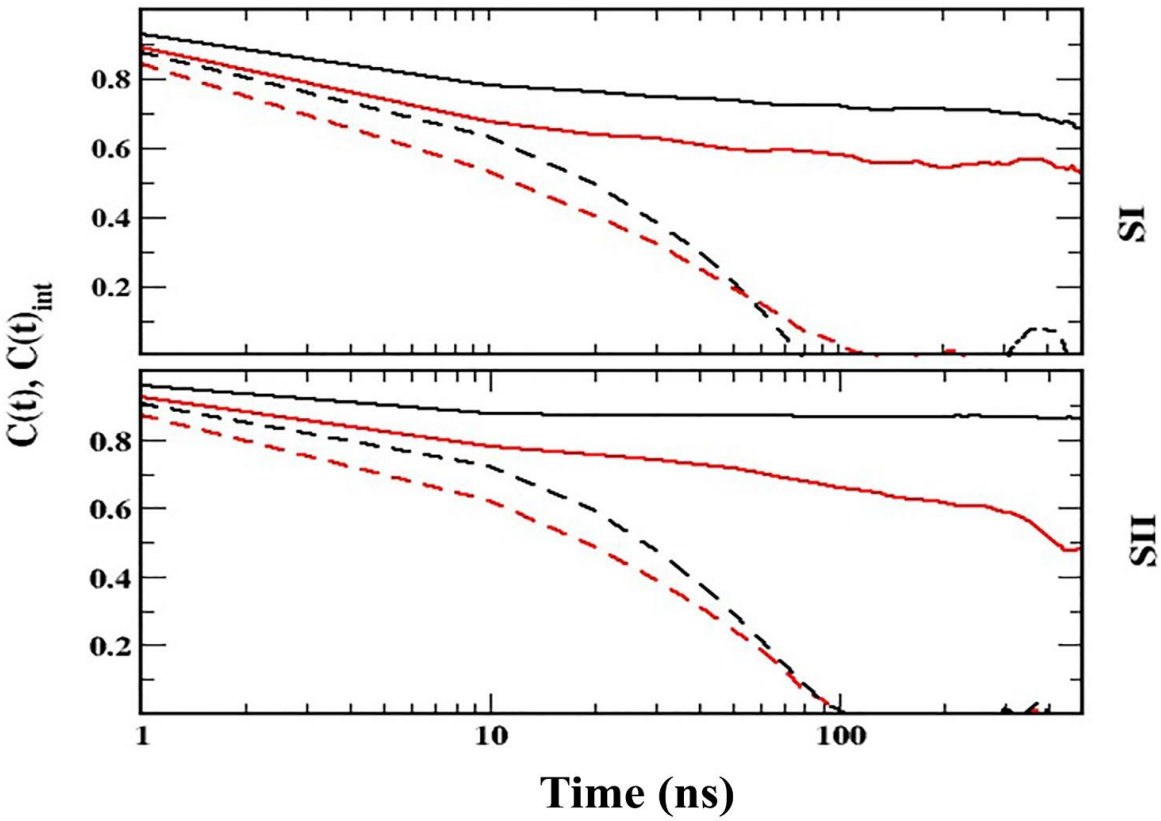

Figure 3.8: Internal and total dynamics correlation functions, the relaxation of the backbone N-H dipoles of the switches regions for both mutants G12D-K-Ras (black) and the oxidized variant (red), with the internal (solid line) and total (dashed line) autocorrelation function.

Therefore, we can conclude from these results that the oxidized G12D-K-Ras sample the more dynamic state. According to Lu et al's study [100] on the mutant H-Ras, the SI region undergoes minor conformational changes in the intermediate state of the A59G-H-Ras mutant, whereas the SII region goes through an extensive transitions between the two structures, indicating the switch region's high flexibility. Therefore, it is reasonable to hypothesize that the oxidized G12D-K-Ras mutant structure represents the intermediate between the conformational transition from Ras-GTP to Ras-GDP or the GTP- conformational state I.

3.6 Sodium-Ion Interaction with GTP in the Active Site

It has been noticed by observation of the MD trajectory of both mutants that a sodium ion is tended to be detained near the GTP: $O\beta_3$ and the negatively charged amino acids in the binding site. So we utilized the ion-GTP: $O\beta_3$ distance as a parameter to quantify the ions' locations in the binding site. Figure (3.9) A close examination of the ions' motion clearly shows that when they get close to the binding site, they tend to stay there. As we can see the ion is drawn to the attracter site within a few nanoseconds of the production phase, but after a while the attracted Na^+ is exchanged many times, 3 in G12D, and 5 in SOH-G12D, suggesting that the binding is somewhat weak and variable. in Kla¨hn study of QM/MM simulations performed on Ras protein, a Na^+ ion diffused rapidly toward the active site, this configuration appears to be stable because the cation compensates for the active site net charge (GTP, Mg^{+2}, and $Lys16^+$), which is -1 e. More broadly, there is a high probability that a Na^+ cation would be found here [104].

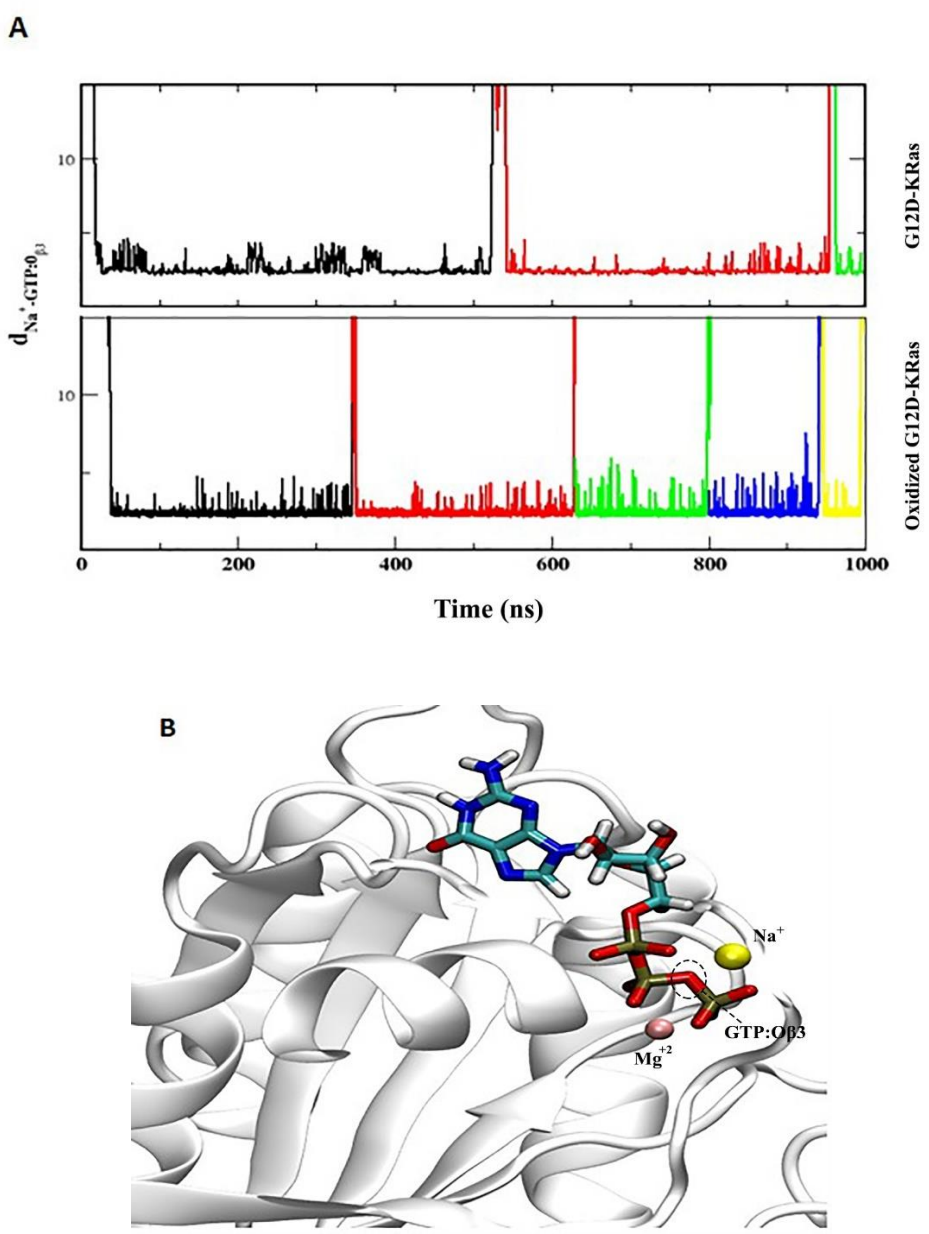

Figure 3.9: Long residence sodium ion in the binding site. A) show the distance between the sodium ion and the GTP: Oβ3, the color change indicates exchange in the ion. B) A snapshot from the simulation showing the sodium ion in the binding site interacting with the GTP.

In previous studies, suggesting that binding of a metal ion at the active site causes Y32 displacement away from the GTP, which contributes to a conformational changes in the SI region affecting the interactions with effectors [5], [15], [105].

Chapter 4: Conclusions

K-Ras is the most commonly mutated oncogene in human cancers, many studies show that cancers are caused by misregulation of Ras redox signaling. It's found that oxidation of Cys118 in Ras proteins provides a mechanism for rapidly and reversibly altering protein function**s** hence, protein activity, stability, and localization, as well as protein-protein interaction, are affected [4], [13], [106].

This investigation reports the results of all-atom MD simulations conducted on the GTP-bound states of the catalytic domain of G12D-K-Ras and its oxidized variant SOH-G12D in an aqueous solvent. This study aimed to investigate the effects of Cys118 sulfinylation on the structure and conformational dynamics of G12D-K-Ras, and thereby obtain further insight into the impact of such dynamics on their functional activities. To achieve this purpose, we performed detailed analyses of intra-protein, protein-GTP, and protein-solvent interactions. Also, we used many measures and techniques to analyze the conformational fluctuations such as RMSD, RMSF, relaxation times, NMR order parameter, PCA, and other techniques. The results highlighted significant structural differences in the two mutants. In particular, there are distinct dynamics and interaction patterns in the GTP binding site and a number of loops distant from it.

The mutants show differences in the conformational dynamics and equilibrium structures, where perturbations and conformation rearrangements to G12D-K-Ras structure have occurred, that make the secondary structure of the oxidized protein experienced larger conformational changes. Significant effects were also observed in the dynamical behavior of the two switch regions, particularly in SII, which induces high flexibility at these regions and some other loops in the oxidized protein The high flexibility of SII in the oxidized variant has an important role in Ras conformational switching between GDP (off) and GTP (on) states, as

well as hydrolysis impairment in Ras mutations at the switches regions or P-loop [107], [108], which influences the nucleotide exchange rate and intrinsic GTPase activity and effector binding. As it's found that both the GDP/GTP or GTP state I/state II transitions are caused by the highly flexible nature of the two switch regions, particularly SII. These experimental results were further supported by a larger region of phase space in the PCA analysis of the oxidized mutant.

From the conformational sampling of G12D and the oxidized variant SOH-G12D, the switches remain relatively closed and hence both mutants remain in the intermediate GTP state I, however, there has been a shift in the conformational state, comparable to the known experimental structures of GTP-analogue bound Ras.

The analysis also has provided insights into the interaction of sodium ions from the solvent with the GTP-phosphate and neighboring residues in the mutants, resulting in conformational changes in SI region, which affects the interactions with effectors. Also we discovered from the comparison of the Cys118 dihedral in G12D and its oxidized counterpart trajectories, a significant transition in the orientation of the side chain of Cys118 due to the sulfenylation of the Cys118-thiol group resulting in perturbation in Phe28 – Cys118 interaction.

These conformational analyses provided clues about dynamic behavior change in the two switches and other regions which may change the preference of Ras for different effectors and disrupt the conformational states [21], [109]. That may lead us to speculate that the oxidized G12D-K-Ras mutant structure could be similar of the intermediate in the conformational transmission from Ras-GTP to Ras-GDP, or the GTP- conformational state I.

According to a review of the literature, this is the first case in which MD simulations was used to investigate the molecular behavior of oxidized G12D-K-Ras. The findings presented here lay the groundwork for future experimental and computational investigations into the

functional consequences of SOH-G12D dynamics, which will provide important insight into approaches to more effectively target oncogenic K-Ras.

References

[1] Mitin, N., K.L. Rossman, and C.J. Der. 2005. Signaling Interplay in Ras Superfamily Function. *Current Biology*. 15:R563–R574.

[2] Wennerberg, K. 2005. The Ras superfamily at a glance. *Journal of Cell Science*. 118:843–846.

[3] Rebollo, A., and C. Martínez-A. 1999. Ras Proteins: Recent Advances and New Functions. *Blood*. 94:2971–2980.

[4] Messina, S., G. De Simone, and P. Ascenzi. 2019. Cysteine-based regulation of redox-sensitive Ras small GTPases. *Redox Biol*. 26:101282.

[5] Sayyed-Ahmad, A., P. Prakash, and A.A. Gorfe. 2017. Distinct dynamics and interaction patterns in H- and K-Ras oncogenic P-loop mutants: Interaction Patterns in Ras Oncogenic Mutants. *Proteins*. 85:1618–1632.

[6] Omerovic, J., A.J. Laude, and I.A. Prior. 2007. Ras proteins: paradigms for compartmentalised and isoform-specific signalling. *Cell. Mol. Life Sci.* 64:2575–2589.

[7] Prior, I.A., P.D. Lewis, and C. Mattos. 2012. A Comprehensive Survey of Ras Mutations in Cancer. *Cancer Res*. 72:2457–2467.

[8] Baussand, J., and J. Kleinjung. 2013. Specific Conformational States of Ras GTPase upon Effector Binding. *J Chem Theory Comput*. 9:738–749.

[9] Ferro, E., L. Goitre, S.F. Retta, and L. Trabalzini. 2012. The Interplay between ROS and Ras GTPases: Physiological and Pathological Implications. *J Signal Transduct*. 2012.

[10] Gurung, A., and A. Bhattacharjee. 2015. Significance of Ras Signaling in Cancer and Strategies for its Control. .

[11] Prakash, P., A. Sayyed-Ahmad, and A.A. Gorfe. 2012. The Role of Conserved Waters in Conformational Transitions of Q61H K-ras. *PLoS Comput Biol*. 8.

[12] Adachi, T., D.R. Pimentel, T. Heibeck, X. Hou, Y.J. Lee, B. Jiang, Y. Ido, and R.A. Cohen. 2004. *S* -Glutathiolation of Ras Mediates Redox-sensitive Signaling by Angiotensin II in Vascular Smooth Muscle Cells. *J. Biol. Chem.* 279:29857–29862.

[13] Heo, J. 2011. Redox Control of GTPases: From Molecular Mechanisms to Functional Significance in Health and Disease. *Antioxidants & Redox Signaling*. 14:689–724.

[14] Pantsar, T. 2020. The current understanding of KRAS protein structure and dynamics. *Computational and Structural Biotechnology Journal*. 18:189–198.

[15] Khaled, M., A. Gorfe, and A. Sayyed-Ahmad. 2019. Conformational and Dynamical Effects of Tyr32 Phosphorylation in K-Ras: Molecular Dynamics Simulation and Markov State Models Analysis. *J. Phys. Chem. B*. 123:7667–7675.

[16] Cox, A.D., and C.J. Der. 2002. Ras Family Signaling: Therapeutic Targeting. *Cancer Biology & Therapy*. 1:599–606.

[17] Mo, S.P., J.M. Coulson, and I.A. Prior. 2018. RAS variant signalling. *Biochem Soc Trans*. 46:1325–1332.

[18] Jančík, S., J. Drábek, D. Radzioch, and M. Hajdúch. 2010. Clinical Relevance of KRAS in Human Cancers. *Journal of Biomedicine and Biotechnology*. 2010:e150960.

[19] Barbacid, M. 1987. ras GENES. *Annual Review of Biochemistry*. 56:779–827.

[20] Barbacid, M. 1990. ras oncogenes: their role in neoplasia. *Eur J Clin Invest*. 20:225–235.

[21] Chen, C.-C., T.-K. Er, Y.-Y. Liu, J.-K. Hwang, M.J. Barrio, M. Rodrigo, E. Garcia-Toro, and M. Herreros-Villanueva. 2013. Computational analysis of KRAS mutations: implications for different effects on the KRAS p.G12D and p.G13D mutations. *PLoS One*. 8:e55793.

[22] Clayton, S.J., F.M. Scott, J. Walker, K. Callaghan, K. Haque, T. Liloglou, G. Xinarianos, S. Shawcross, P. Ceuppens, J.K. Field, and J.C. Fox. 2000. K-ras Point Mutation Detection in Lung Cancer: Comparison of Two Approaches to Somatic Mutation Detection Using ARMS Allele-specific Amplification. *Clinical Chemistry*. 46:1929–1938.

[23] Scheffzek, K., M.R. Ahmadian, W. Kabsch, L. Wiesmüller, A. Lautwein, F. Schmitz, and A. Wittinghofer. 1997. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*. 277:333–338.

[24] Lander, H.M., D.P. Hajjar, B.L. Hempstead, U.A. Mirza, B.T. Chait, S. Campbell, and L.A. Quilliam. 1997. A molecular redox switch on p21(ras). Structural basis for the nitric oxide-p21(ras) interaction. *J Biol Chem*. 272:4323–4326.

[25] Kamata, H., and H. Hirata. 1999. Redox regulation of cellular signalling. *Cell Signal*. 11:1–14.

[26] Foo, C.H.J., and S. Pervaiz. 2018. gRASping the redox lever to modulate cancer cell fate signaling. *Redox Biology*. 101094.

[27] Alcock, L.J., M.V. Perkins, and J.M. Chalker. 2018. Chemical methods for mapping cysteine oxidation. *Chem. Soc. Rev.* 47:231–268.

[28] Hobbs, G.A., B. Zhou, A.D. Cox, and S.L. Campbell. 2014. Rho GTPases, oxidation, and cell redox control. *Small GTPases*. 5.

[29] Ray, P.D., B.-W. Huang, and Y. Tsuji. 2012. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling*. 24:981–990.

[30] Grek, C.L., J. Zhang, Y. Manevich, D.M. Townsend, and K.D. Tew. 2013. Causes and Consequences of Cysteine S-Glutathionylation. *J. Biol. Chem.* 288:26497–26504.

[31] Wall, S.B., J.-Y. Oh, A.R. Diers, and A. Landar. 2012. Oxidative Modification of Proteins: An Emerging Mechanism of Cell Signaling. *Front. Physiol.* 3.

[32] Paulsen, C.E., and K.S. Carroll. 2013. Cysteine-Mediated Redox Signaling: Chemistry, Biology, and Tools for Discovery. *Chem. Rev.* 113:4633–4679.

[33] Poole, L.B., P.A. Karplus, and A. Claiborne. 2004. Protein Sulfenic Acids in Redox Signaling. *Annual Review of Pharmacology and Toxicology*. 44:325–347.

[34] England, K., and T.G. Cotter. 2005. Direct oxidative modifications of signalling proteins in mammalian cells and their effects on apoptosis. *Redox Report*. 10:237–245.

[35] Forman, H.J., J.M. Fukuto, and M. Torres. 2004. Redox signaling: thiol chemistry defines which reactive oxygen and nitrogen species can act as second messengers. *American Journal of Physiology-Cell Physiology*. 287:C246–C256.

[36] Armeni, T., L. Ercolani, L. Urbanelli, A. Magini, F. Magherini, A. Pugnaloni, F. Piva, A. Modesti, C. Emiliani, and G. Principato. 2012. Cellular Redox Imbalance and Changes of Protein S-glutathionylation Patterns Are Associated with Senescence Induced by Oncogenic H-Ras. *PLOS ONE*. 7:e52151.

[37] Biswas, S., A.S. Chida, and I. Rahman. 2006. Redox modifications of protein–thiols: Emerging roles in cell signaling. *Biochemical Pharmacology*. 71:551–564.

[38] Lander, H.M., J.S. Ogiste, K.K. Teng, and A. Novogrodsky. 1995. p21ras as a common signaling target of reactive free radicals and cellular redox stress. *J Biol Chem*. 270:21195–21198.

[39] Heo, J., and S.L. Campbell. 2004. Mechanism of p21 ^{Ras} S-Nitrosylation and Kinetics of Nitric Oxide-Mediated Guanine Nucleotide Exchange †. *Biochemistry*. 43:2314–2322.

[40] Hobbs, G.A., M.G. Bonini, H.P. Gunawardena, X. Chen, and S.L. Campbell. 2013. Glutathiolated Ras: Characterization and implications for Ras activation. *Free Radical Biology and Medicine*. 57:221–229.

[41] Satomi, Y., P. Bu, M. Okuda, H. Tokuda, and H. Nishino. 2003. H-ras mutations at codon 61 or 13 in tumors initiated with a NO donor in mouse skin. *Cancer Letters*. 196:17–22.

[42] Huang, L., and C.M. Counter. 2015. Reduced HRAS G12V-Driven Tumorigenesis of Cell Lines Expressing KRAS C118S. *PLoS One*. 10:e0123918.

[43] Huang, L., J. Carney, D.M. Cardona, and C.M. Counter. 2014. Decreased tumorigenesis in mice with a Kras point mutation at C118. *Nat Commun*. 5:5410.

[44] Weinberg, F., N. Ramnath, and D. Nagrath. 2019. Reactive Oxygen Species in the Tumor Microenvironment: An Overview. *Cancers (Basel)*. 11.

[45] Ha, J.-H., and S.N. Loh. 2012. Protein Conformational Switches: From Nature to Design. *Chemistry*. 18:7984–7999.

[46] Haspel, N., M. Moll, M.L. Baker, W. Chiu, and L.E. Kavraki. 2010. Tracing conformational changes in proteins. *BMC Struct Biol*. 10:S1.

[47] Orellana, L. 2019. Large-Scale Conformational Changes and Protein Function: Breaking the in silico Barrier. *Front. Mol. Biosci.* 6.

[48] Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem*. 26:1781–1802.

[49] MacKerell, A.D., M. Feig, and C.L. Brooks. 2004. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc*. 126:698–699.

[50] Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. 14:33–38.

[51] Grant, B.J., A.P.C. Rodrigues, K.M. ElSawy, J.A. McCammon, and L.S.D. Caves. 2006. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*. 22:2695–2696.

[52] Turner, P. 2005. XMGRACE, Version 5.1. 19. *Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR*.

[53] Attig, N., K. Binder, H. Grubmüller, K. Kremer, P.G. Betriebe, and F. Jülich. 2004. Poster Abstracts. .

[54] Albano, J.M.R., E. de Paula, and M. Pickholz. 2018. Molecular Dynamics Simulations to Study Drug Delivery Systems. *Molecular Dynamics*.

[55] Durrant, J.D., and J.A. McCammon. 2011. Molecular dynamics simulations and drug discovery. *BMC Biology*. 9:71.

[56] Hollingsworth, S.A., and R.O. Dror. 2018. Molecular Dynamics Simulation for All. *Neuron*. 99:1129–1143.

[57] Karplus, M., and J.A. McCammon. 2002. Molecular dynamics simulations of biomolecules. *Nat. Struct Biol.* 9:646–652.

[58] González, M.A. 2011. Force fields and molecular dynamics simulations. *JDN*. 12:169–200.

[59] Patodia, S. 2014. Molecular Dynamics Simulation of Proteins: A Brief Overview. *2161-0398*. 4.

[60] Brooks, B.R., C.L. Brooks, A.D. MacKerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M.

Karplus. 2009. CHARMM: The Biomolecular Simulation Program. *J Comput Chem*. 30:1545–1614.

[61] Galindo-Murillo, R., J.C. Robertson, M. Zgarbová, J. Šponer, M. Otyepka, P. Jurečka, and T.E. Cheatham. 2016. Assessing the Current State of Amber Force Field Modifications for DNA. *J Chem Theory Comput*. 12:4114–4127.

[62] Bernardes, C.E.S., and A. Joseph. 2015. Evaluation of the OPLS-AA Force Field for the Study of Structural and Energetic Aspects of Molecular Organic Crystals. *J. Phys. Chem. A*. 119:3023–3034.

[63] Horta, B.A.C., P.T. Merz, P.F.J. Fuchs, J. Dolenc, S. Riniker, and P.H. Hünenberger. 2016. A GROMOS-Compatible Force Field for Small Organic Molecules in the Condensed Phase: The 2016H66 Parameter Set. *J. Chem. Theory Comput.* 12:3825–3850.

[64] Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. MacKerell. 2010. CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 31:671–690.

[65] Jo, S., T. Kim, V.G. Iyer, and W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem*. 29:1859–1865.

[66] Heppner, D.E., C.M. Dustin, C. Liao, M. Hristova, C. Veith, A.C. Little, B.A. Ahlers, S.L. White, B. Deng, Y.-W. Lam, J. Li, and A. van der Vliet. 2018. Direct cysteine sulfenylation drives activation of the Src kinase. *Nat Commun*. 9:4522.

[67] Olsson, M.H.M., C.R. Søndergaard, M. Rostkowski, and J.H. Jensen. 2011. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* 7:525–537.

[68] Ryckaert, J.-P., G. Ciccotti, and H.J.C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*. 23:327–341.

[69] Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*. 98:10089–10092.

[70] Pastor, R.W. 1994. Techniques and Applications of Langevin Dynamics Simulations. In: Luckhurst GR, CA Veracini, editors. The Molecular Dynamics of Liquid Crystals. Dordrecht: Springer Netherlands. pp. 85–138.

[71] Feller, S.E., Y. Zhang, R.W. Pastor, and B.R. Brooks. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103:4613–4621.

[72] Bell, E.W., and Y. Zhang. 2019. DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. *Journal of Cheminformatics*. 11:40.

[73] Maiorov, V., and G. Crippen. 1994. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *Journal of molecular biology*.

[74] Haider, S., G.N. Parkinson, and S. Neidle. 2008. Molecular Dynamics and Principal Components Analysis of Human Telomeric Quadruplex Multimers. *Biophysical Journal*. 95:296–311.

[75] Kufareva, I., and R. Abagyan. 2012. Methods of protein structure comparison. *Methods Mol Biol*. 857:231–257.

[76] Martínez, L. 2015. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS One*. 10:e0119264.

[77] Al Qaraghuli, M.M., K. Kubiak-Ossowska, and P.A. Mulheran. 2018. Thinking outside the Laboratory: Analyses of Antibody Structure and Dynamics within Different Solvent Environments in Molecular Dynamics (MD) Simulations. *Antibodies (Basel)*. 7:21.

[78] Palmer, A.G. 2004. NMR characterization of the dynamics of biomacromolecules. *Chem Rev*. 104:3623–3640.

[79] Trbovic, N., B. Kim, R.A. Friesner, and A.G. Palmer. 2008. Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. *Proteins*. 71:684–694.

[80] Sharp, K.A., E. O'Brien, V. Kasinath, and A.J. Wand. 2015. On the relationship between NMR-derived amide order parameters and protein backbone entropy changes. *Proteins*. 83:922–930.

[81] Gu, Y., D.-W. Li, and R. Brüschweiler. 2014. NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. *J Chem Theory Comput*. 10:2599–2607.

[82] Dubyna, V.M., D. Kovalskyy, O.S. Ivanova, and A. Kornelyuk. 2006. The improvement of the algorithm for order parameter calculation (S2) from molecular dynamics simulation using the correlation motion function. *Biophysical chemistry*. 123:25–8.

[83] Perilla, J.R., and T.B. Woolf. 2012. Towards the prediction of order parameters from molecular dynamics simulations in proteins. *J. Chem. Phys.* 136:164101.

[84] Eriksson, M.A., H. Berglund, T. Härd, and L. Nilsson. 1993. A comparison of 15N NMR relaxation measurements with a molecular dynamics simulation: backbone dynamics of the glucocorticoid receptor DNA-binding domain. *Proteins*. 17:375–390.

[85] Pálfy, G., D.K. Menyhárd, and A. Perczel. 2020. Dynamically encoded reactivity of Ras enzymes: opening new frontiers for drug discovery. *Cancer Metastasis Rev*. 39:1075–1089.

[86] Chandrasekhar, I., G.M. Clore, A. Szabo, A.M. Gronenborn, and B.R. Brooks. 1992. A 500 ps molecular dynamics simulation study of interleukin-1 beta in water. Correlation with nuclear magnetic resonance spectroscopy and crystallography. *J Mol Biol*. 226:239–250.

[87] Hoffmann, F., F.A.A. Mulder, and L.V. Schäfer. 2020. Predicting NMR relaxation of proteins from molecular dynamics simulations with accurate methyl rotation barriers. *J. Chem. Phys.* 152:084102.

[88] Balsera, M., W. Wriggers, Y. Oono, and K. Schulten. 1996. Principal Component Analysis and Long Time Protein Dynamics. .

[89] Altis, A., P.H. Nguyen, R. Hegger, and G. Stock. 2007. Dihedral angle principal component analysis of molecular dynamics simulations. *J. Chem. Phys.* 126:244111.

[90] Wolf, A., and K.N. Kirschner. 2013. Principal component and clustering analysis on molecular dynamics data of the ribosomal L11·23S subdomain. *J Mol Model*. 19:539–549.

[91] Sittel, F., A. Jain, and G. Stock. 2014. Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates. *J. Chem. Phys.* 141:014111.

[92] David, C., and D. Jacobs. 2014. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods in molecular biology (Clifton, N.J.)*. 1084:193–226.

[93] Amadei, A., A.B. Linssen, B.L. de Groot, D.M. van Aalten, and H.J. Berendsen. 1996. An efficient method for sampling the essential subspace of proteins. *J Biomol Struct Dyn*. 13:615–625.

[94] Skjaerven, L., N. Reuter, and A. Martinez. 2011. Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules: a computational approach. *Future Med Chem*. 3:2079–2100.

[95] Shima, F., Y. Ijiri, S. Muraoka, J. Liao, M. Ye, M. Araki, K. Matsumoto, N. Yamamoto, T. Sugimoto, Y. Yoshikawa, T. Kumasaka, M. Yamamoto, A. Tamura, and T. Kataoka. 2010. Structural Basis for Conformational Dynamics of GTP-bound Ras Protein. *J Biol Chem*. 285:22696–22705.

[96] Dash, R., M.C. Ali, N. Dash, M.A.K. Azad, S.M.Z. Hosen, M.A. Hannan, and I.S. Moon. 2019. Structural and Dynamic Characterizations Highlight the Deleterious Role of SULT1A1 R213H Polymorphism in Substrate Binding. *Int J Mol Sci*. 20:E6256.

[97] Pacold, M.E., S. Suire, O. Perisic, S. Lara-Gonzalez, C.T. Davis, E.H. Walker, P.T. Hawkins, L. Stephens, J.F. Eccleston, and R.L. Williams. 2000. Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell*. 103:931–943.

[98] Prakash, P., and A.A. Gorfe. 2013. Lessons from computer simulations of Ras proteins in solution and in membrane. *Biochim Biophys Acta*. 1830:10.1016/j.bbagen.2013.07.024.

[99] Kobayashi, C., and S. Saito. 2010. Relation between the Conformational Heterogeneity and Reaction Cycle of Ras: Molecular Simulation of Ras. *Biophys J*. 99:3726–3734.

[100] Lu, S., H. Jang, S. Muratcioglu, A. Gursoy, O. Keskin, R. Nussinov, and J. Zhang. 2016. Ras Conformational Ensembles, Allostery, and Signaling. *Chem Rev*. 116:6607–6665.

[101] Lu, J., R.A. Harrison, L. Li, M. Zeng, S. Gondi, D. Scott, N.S. Gray, J.R. Engen, and K.D. Westover. 2017. KRAS G12C Drug Development: Discrimination between Switch II Pocket Configurations Using Hydrogen/Deuterium-Exchange Mass Spectrometry. *Structure*. 25:1442-1448.e3.

[102] Spoerner, M., C. Hozsa, J.A. Poetzl, K. Reiss, P. Ganser, M. Geyer, and H.R. Kalbitzer. 2010. Conformational states of human rat sarcoma (Ras) protein complexed with its natural ligand GTP and their role for effector interaction and GTP hydrolysis. *J Biol Chem*. 285:39768–39778.

[103] Ollila, O.H.S., H.A. Heikkinen, and H. Iwaï. 2018. Rotational Dynamics of Proteins from Spin Relaxation Times and Molecular Dynamics Simulations. *J Phys Chem B*. 122:6559–6569.

[104] Klähn, M., J. Schlitter, and K. Gerwert. 2005. Theoretical IR Spectroscopy Based on QM/MM Calculations Provides Changes in Charge Distribution, Bond Lengths, and Bond Angles of the GTP Ligand Induced by the Ras-Protein. *Biophysical Journal*. 88:3829–3844.

[105] Miao, W., L. Eichelberger, L. Baker, and M.S. Marshall. 1996. p120 Ras GTPase-activating Protein Interacts with Ras-GTP through Specific Conserved Residues *. *Journal of Biological Chemistry*. 271:15322–15329.

[106] Tong, L.A., A.M. de Vos, M.V. Milburn, and S.H. Kim. 1991. Crystal structures at 2.2 A resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GDP. *J Mol Biol*. 217:503–516.

[107] Visscher, M., M.R. Arkin, and T.B. Dansen. 2016. Covalent targeting of acquired cysteines in cancer. *Current Opinion in Chemical Biology*. 30:61–67.

[108] Spoerner, M., C. Herrmann, I.R. Vetter, H.R. Kalbitzer, and A. Wittinghofer. 2001. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc Natl Acad Sci U S A*. 98:4944–4949.

[109] Spoerner, M., A. Wittinghofer, and H.R. Kalbitzer. 2004. Perturbation of the conformational equilibria in Ras by selective mutations as studied by 31P NMR spectroscopy. *FEBS Lett*. 578:305–310.

Appendices

Appendix I: Force Field Parameters for Cysteine sulfenic acid (SOH)

>>>Parameters included in CHARMM forcefield parameter file

BONDS
!
! V(bond) = Kb(b - b0)2**
! Kb: kcal/mole/A2**
! b0: A
!
! atom type Kb b0
SO1 OH1 200.000 1.7090 !
SO1 CT2 198.000 1.8230 !

ANGLES
!
! V(angle) = Ktheta (Theta - Theta0) **2
! Ktheta: kcal/mole/rad**2
! Theta0: degrees
!
! atom types Ktheta Theta0
CT1 CT2 SO1 50.000 115.3000 !
SO1 OH1 H 50.000 110.1000 !
CT2 SO1 OH1 50.000 99.0000 !
SO1 CT2 HA2 46.100 107.0000 !

DIHEDRALS
!
! V (dihedral) = Kchi (1 + cos (n (chi) - delta))
! Kchi: kcal/mole
! n: multiplicity
! delta: degrees
!
! atom types Kchi n delta
CT1 CT2 SO1 OH1 0.2000 3 0.00 !
NH1 CT1 CT2 SO1 0.2000 3 0.00 !
H C CT2 SO1 0.2000 3 0.00 !
C CT2 SO1 OH1 0.2000 3 0.00 !
HA2 CT2 SO1 OH1 0.2000 3 0.00 !
H OH1 SO1 CT2 1.1000 2 0.00 !
