

## Article

# Markov State Models and Molecular Dynamics Simulations Provide Understanding of the Nucleotide-Dependent Dimerization-Based Activation of LRRK2 ROC Domain

Xinyi Li <sup>1,2,†</sup>, Zengxin Qi <sup>3,4,5,†</sup>, Duan Ni <sup>2</sup> , Shaoyong Lu <sup>2,\*</sup>, Liang Chen <sup>3,4,5,\*</sup> and Xiangyu Chen <sup>1,\*</sup>

<sup>1</sup> School of Medical Laboratory, Weifang Medical University, Weifang 261053, China; Ashley\_Li@sjtu.edu.cn

<sup>2</sup> Medicinal Chemistry and Bioinformatics Center, School of Medicine, Shanghai Jiao Tong University, Shanghai 200025, China; niduan11@sjtu.edu.cn

<sup>3</sup> Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai 200040, China; qizengxin@huashan.org.cn

<sup>4</sup> Shanghai Key Laboratory of Brain Function Restoration and Neural Regeneration, Huashan Hospital, Fudan University, Shanghai 200040, China

<sup>5</sup> State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, School of Basic Medical Sciences and Institutes of Brain Science, Fudan University, Shanghai 200433, China

\* Correspondence: lushaoyong@sjtu.edu.cn (S.L.); chenlianghs@126.com (L.C.); xiangyu0711@wfm.edu.cn (X.C.)

† These authors contributed equally to this work.



**Citation:** Li, X.; Qi, Z.; Ni, D.; Lu, S.; Chen, L.; Chen, X. Markov State Models and Molecular Dynamics Simulations Provide Understanding of the Nucleotide-Dependent Dimerization-Based Activation of LRRK2 ROC Domain. *Molecules* **2021**, *26*, 5647. <https://doi.org/10.3390/molecules26185647>

Academic Editors: Stacey Wetmore, Angelo Facchiano and Anna Marabotti

Received: 22 August 2021

Accepted: 14 September 2021

Published: 17 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Mutations in leucine-rich repeat kinase 2 (LRRK2) are recognized as the most frequent cause of Parkinson's disease (PD). As a multidomain ROCO protein, LRRK2 is characterized by the presence of both a Ras-of-complex (ROC) GTPase domain and a kinase domain connected through the C-terminal of an ROC domain (COR). The bienzymatic ROC–COR–kinase catalytic triad indicated the potential role of GTPase domain in regulating kinase activity. However, as a functional GTPase, the detailed intrinsic regulation of the ROC activation cycle remains poorly understood. Here, combining extensive molecular dynamics simulations and Markov state models, we disclosed the dynamic structural rearrangement of ROC's homodimer during nucleotide turnover. Our study revealed the coupling between dimerization extent and nucleotide-binding state, indicating a nucleotide-dependent dimerization-based activation scheme adopted by ROC GTPase. Furthermore, inspired by the well-known R1441C/G/H PD-relevant mutations within the ROC domain, we illuminated the potential allosteric molecular mechanism for its pathogenetic effects through enabling faster interconversion between inactive and active states, thus trapping ROC in a prolonged activated state, while the implicated allostery could provide further guidance for identification of regulatory allosteric pockets on the ROC complex. Our investigations illuminated the thermodynamics and kinetics of ROC homodimer during nucleotide-dependent activation for the first time and provided guidance for further exploiting ROC as therapeutic targets for controlling LRRK2 functionality in PD treatment.

**Keywords:** Parkinson's disease; leucine-rich repeat kinase 2 (LRRK2); Ras-of-complex GTPase domain; molecular dynamics (MD) simulations; Markov state models; network analysis

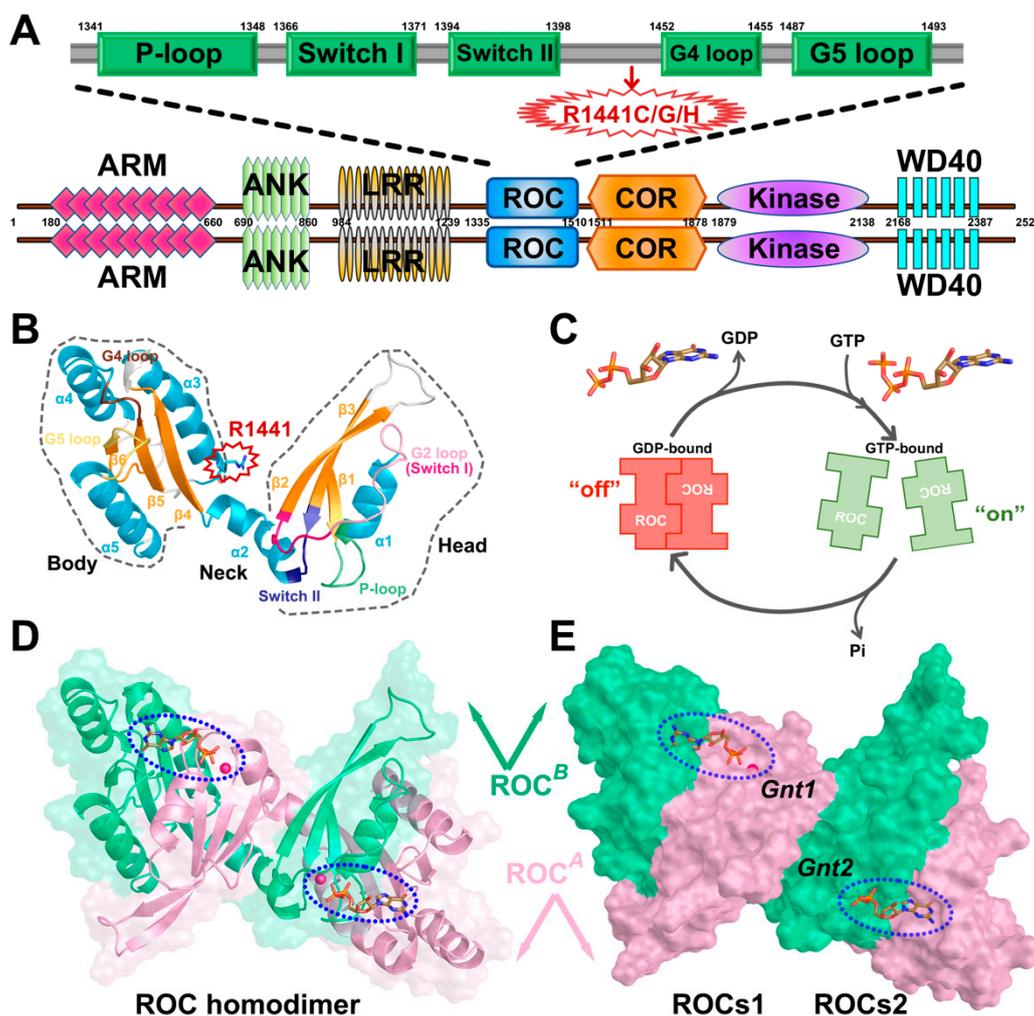
## 1. Introduction

The rapidly ageing population has imposed challenging public health crises in dealing with age-associated neurodegenerative afflictions with effective treatments [1–3]. As one of the most prevalent neurodegenerative disorders, Parkinson's disease (PD) features symptoms including muscular rigidity, resting tremor, and movement slowness [4,5]. Pathologically, accumulating evidence has suggested that progressive degeneration of dopaminergic (DA) neurons and the presence of abnormal intraneuronal cytoplasmic inclusions (termed Lewy bodies) rich in  $\alpha$ -synuclein are hallmarks in PD pathogenesis [6–8]. Although intense

efforts have been made to understand this desperate disease, the etiology of PD remains unresolved, while both environmental and genetic factors are suggested to play a part [9–11]. To date, missense mutations in leucine-rich repeat Kinase 2 (LRRK2), the gene product of Parkinson's recognized risk loci PARK8, have been recognized as the most frequent cause of PD among all relevant pathogenetic genes identified [12–14]. LRRK2 belongs to the superfamily of ROCO proteins, a unique feature of which is the presence of a Ras-of-complex domain (ROC) in tandem with a C-terminal of a ROC domain (COR), immediately followed by a tyrosine kinase-like protein kinase domain (kinase) [15,16]. Meanwhile, additional protein–protein interaction domains are also found in the N-terminus (armadillo: ARM, ankyrin: ANK, and leucine-rich repeats: LRR) and C-terminus (WD40 repeats) of the central ROC–COR–kinase catalytic triad of LRRK2 [17–19] (Figure 1A). ROC contains five  $\alpha$  helices and six  $\beta$  strands connected through loops, and can be roughly divided into head ( $\beta$ 1,  $\alpha$ 1,  $\beta$ 2, and  $\beta$ 3), neck ( $\alpha$ 2), and body ( $\beta$ 4,  $\alpha$ 3,  $\beta$ 5,  $\alpha$ 4,  $\beta$ 6, and  $\alpha$ 5) subdomains [20] (Figure 1B). Crystallographic studies have revealed the unique homodimeric structure of ROC, as LRRK2 exists predominantly in the form of dimers in cells and tissues (ROC dimer: ROCs, Figure 1D). With the domain-swapping dimerization scheme, ROCs form two nearly symmetric nucleotide-binding sites (referred to as *Gnt1* and *Gnt2*) with structural motifs contributed from both monomers (referred to as ROC<sup>A</sup> and ROC<sup>B</sup>, Figure 1E). Moreover, subdomains of each ROC monomer can further pair with another and form compact functional units with corresponding nucleotide-binding sites (ROCs1 and ROCs2, Figure 1E), which resembles the classic scheme adopted by small GTPase [21–23]. Of note, the bienzymatic architecture of LRRK2 implies that the GTPase and kinase activity might be coupled via an intramolecular mechanism with GTPase activity serving to regulate kinase activity [24–26], and this notion is further supported by the observation that disease-causing missense mutations within the ROC domain, the famous R1441C/G/H point mutations, for instance [27–29], usually result in up-regulated kinase activity [14,30]. However, as a functional GTPase, the detailed mode of action of ROC remains elusive, but studies on LRRK2 homologs from prokaryotes have suggested that LRRK2 ROC is likely to function as a G protein activated by nucleotide-dependent dimerization (GAD) [31], which relies mainly on spontaneous nucleotide turnover and dimerization for the regulation of the activation cycle [32,33] (Figure 1C). Under this circumstance, exploration of the conformational dynamics relevant to the dimerization and nucleotide turnover of LRRK2 ROC is of primary importance for further investigations into its contribution to the overall functional output of LRRK2 under both physiological and pathological conditions.

Nonetheless, with the existing experimental methodologies, it remains challenging to describe the dynamic conformational transition process, especially in the case of binding ligand turnover, which usually results in distinctive structures with different bioactivity [19,34–37]. To obtain a comprehensive insight into the dynamics of ROC homodimer during nucleotide shift, we employed large-scale all-atom molecular dynamics (MD) simulation to observe *in silico* at the atomistic level the intrinsic regulation of GTPase activation cycle of the ROC domain of LRRK2. Furthermore, by integrating our extensive MD simulations with the statistically robust Markov state models (MSMs) for interpreting configuration sampling [38–40], kinetically relevant states of ROCs homodimer as well as their interconversion rates during GDP/GTP exchange are accessible. Based on MSMs, our study provided for the first time the dynamic portrait of ROCs during nucleotide turnover, as characterized by the previously unreported coupling between dimerization and nucleotide-binding state, with the ROCs dimer exhibiting the “open” conformation with greater dimerization extent in the presence of GDP but switching to the “closed” structure with oligomerization tendency while in complex with GTP. Meanwhile, through dynamic network analysis [41], we dissected the intramolecular correlation network mediating such global structural rearrangements. Furthermore, inspired by the disease-causing mutations clustered within ROC domain, we found that such mutants, as exemplified by R1441C/G/H, possess the potential to remarkably accelerate the interconversion between GDP-bound inactive and GTP-bound active states. The global effects

introduced by such single-residue substitution indicated potential allosteric regulation, and we thus probed into the allosteric signaling pathway associated with mutations. We demonstrated that signals associated with R1441C/G/H mutation could propagate toward the nucleotide-binding sites of ROCs, and more importantly, we disclosed the  $\alpha 2$ ,  $\beta 4$  of both ROC monomers as relatively conservative components for transmitting such allosteric signals, providing guidance for the identification of potential allosteric pockets. Our studies supplement for the first time a dynamic picture of the LRRK2 ROC domain during nucleotide turnover, which supported the presumption that LRRK2 adopted a nucleotide-dependent dimerization-activation scheme like that of GADs. In the meantime, we also uncovered the detailed molecular mechanism for the pathogenetic effects of R1441C/G/H mutations, which further led us to the identification of important allosteric signal transducers. The results offer an insightful understanding toward of the GTPase activation cycle of the ROC domain as well as its role in regulating LRRK2 functionality, and such mechanistic insights provide guidance for targeting ROCs for tuning LRRK2 bioactivity for future PD treatment.



**Figure 1.** (A) Domain architecture of human LRRK2 with its respective amino acid positions. ARM: Armadillo repeats region; ANK: Ankyrin repeat region; LRR: leucine-rich repeats; ROC: Ras-of-complex GTPase domain; COR: C-terminal-of-ROC domain; Kinase: protein tyrosine kinase-like domain; WD40: WD40 repeat. The detailed functional partitioning of ROC is depicted above with the guanine nucleotide phosphate-binding loop (P-loop), Switch I and Switch II, and G4 and G5 loop highlighted in green rectangles. The residue position R1441 is highlighted with a red arrow. (B) Cartoon model of ROC monomer roughly divided into head (including  $\beta 1$ ,  $\alpha 1$ ,  $\beta 2$ , and  $\beta 3$ ), neck (including  $\alpha 2$ ) and body (including  $\beta 4$ ,  $\alpha 3$ ,  $\beta 5$ ,  $\alpha 4$ ,  $\beta 6$ , and  $\alpha 5$ ) subdomains. The monomer is colored based on a secondary structure, and the P-loop, Switch I, Switch II, G4

loop, and G5 loop are highlighted in light green, hot pink, purple, brown, and gold, respectively. Residue R1441 is highlighted. (C) Schematic model of ROC dimer–monomer dynamic transition during nucleotide turnover based on GAD theory. (D) Cartoon model of ROCs-GDP, the two monomers are shown in pink ( $\text{ROC}^A$ ) and green ( $\text{ROC}^B$ ), respectively. (E) Molecular surface representation of ROC dimer highlighting the nucleotide-binding pockets (blue-dashed oval). *Gnt1* is accommodated by  $\text{ROC}^A$  head and  $\text{ROC}^B$  body, while *Gnt2* interacts with  $\text{ROC}^A$  body and  $\text{ROC}^B$  head. Unless otherwise specified, the graphs showing ROCs overview in our work will all be presented from this perspective with *Gnt1* residing on the top-left. The presumed compact functional units are denoted as ROCs1 and ROCs2.

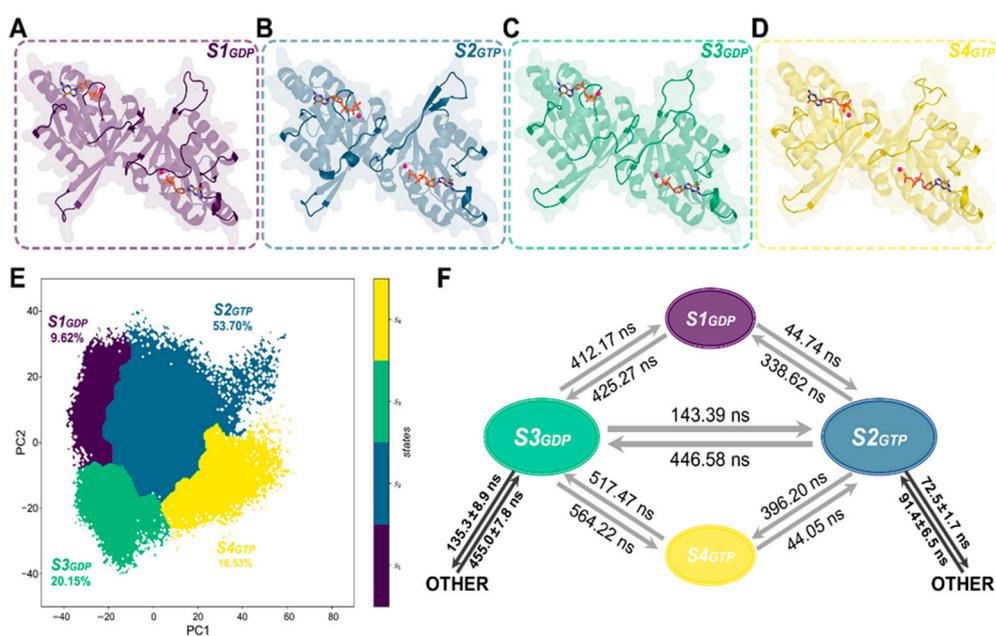
## 2. Results

### 2.1. Construction and Validation of Markov State Models (MSMs)

The solved crystal structure of ROC homodimer (ROCs) in association with either guanosine diphosphate (ROCs-GDP, PDB ID: 2ZEJ [20]) or guanosine triphosphate (ROCs-GTP, artificially built by substituting GDP with GTP in ROCs-GDP complex, detailed in Section 4) were subjected to unbiased large-scale classic MD simulations to dissect the dynamic conformational regulation contributed by guanine nucleotide turnover at an atomistic level. With the refined starting crystal structures, we gathered a total of 6  $\mu\text{s}$  extensive sampling with  $1 \times 3 \mu\text{s}$  independent runs with random initial velocities for both systems, which permitted more complete and thorough mapping of the conformational landscape of ROCs throughout its activation cycle. Both systems reached equilibrium during 1  $\mu\text{s}$  MD simulation in every replica (Section 4, Figure S1). Furthermore, to identify the key components along the activation pathway, we employed the kinetic network MSM embedded in PyEMMA [42]. The simulation trajectories of both ROCs-GDP and ROCs-GTP were first featurized with inter-residue distances to describe the overall ROCs topology, and the resulting dynamics data were subjected to principal component analysis (PCA) from *scikit-learn* for dimensionality reduction to preserve the first two components (PC1 and PC2) with the greatest contribution to kinetic variance [43]. The Markovianity of our model was validated through implied timescales estimations [44,45] (Figure S2A). A lag time of 3 ns was selected for MSM construction, and 100 microstate centers were chosen for the *k-means* clustering algorithm [46]. The obtained microstates were further clustered into four metastable states (also called macrostates) using the PCCA+ algorithm with a passing Chapman-Kolmogorov test [47,48] (Figure S3), obtaining the ROCs-GDP~ROCs-GTP MSM that contains the thermodynamics and thermokinetics of ROCs during activation (Figure 2E).

### 2.2. MSMs Revealed Different Dimerization Extent of ROCs during Nucleotide Turnover

Based on our constructed ROCs-GDP~ROCs-GTP MSMs, the conformational space of ROCs during guanine nucleotide turnover was divided into four metastable structures as shown in Figure 2A. From the distribution of ROCs-GDP and ROCs-GTP on the 2-dimensional plane defined by the two principal components (Figure S4A, Table S1), metastable structures *S1* and *S3* were found to be dominated by the ROCs-GDP conformational ensemble and more likely to represent the “inactive” resting conformation of ROCs under GDP-bound state. In contrast, a greater population of ROCs were found to bind GTP in macrostates *S2* and *S4*, which might denote the active state. To identify key structural features of ROCs under either GDP- or GTP-bound states, we extracted the representative conformations from the four metastables in their more prominent nucleotide-binding state and obtained  $S1_{\text{GDP}}$ ,  $S2_{\text{GTP}}$ ,  $S3_{\text{GDP}}$ , and  $S4_{\text{GTP}}$  for further investigations (Figure 2A–E, Table S1).



**Figure 2.** Representative structures of the four metastable states in ROCs-GDP~ROCs-GTP MSMs: (A)  $S1_{GDP}$  for macrostate  $S1$ , (B)  $S2_{GTP}$  for macrostate  $S2$ , (C)  $S3_{GDP}$  for macrostate  $S3$ , and (D)  $S4_{GTP}$  for macrostate  $S4$ . Structural segments that exhibited significant differences are highlighted in the opaque cartoon, with the remaining being transparent. (E) Distribution of the four metastable states in Rocs-GDP~ROCs-GTP MSMs with the portion of each macrostate. (F) Transition time between the macrostates calculated by mean first passage time (MFPT). OTHER: denoted the transition between  $S2_{GTP}/S3_{GDP}$  and all other remaining metastable states.

A closer insight onto these representatives unveiled that Switch I,  $\beta 2$ - $\beta 3$  linker, G4 loop, and G5 loop were the primary sites for conformational rearrangements (compare Figure 2A–D with Figure 1B). As the representative structure for ROCs-GDP and ROCs-GTP,  $S3_{GDP}$  and  $S2_{GTP}$  exhibited significantly different structures in the above-mentioned motifs, with  $S3_{GDP}$  forming a relatively more compact conformation as the implicated structural segments were posing closer to the main body (Figure 2B,C), and this suggested that the extent to which ROC monomer dimerized may be dissimilar. Indeed, looking into the extensive contact network between the ROC monomers proved that either the stability or the strength of the interactions formed between monomers to facilitate dimerization weakened prominently upon substituting GDP with GTP (Table S2). In the meantime, the interfacing surface area between monomers also decreased (Table S3), further weakening the intramolecular interactions and destabilizing dimerization. Moreover, by calculating the binding free energy between the ROC monomers with Molecular Mechanics/Poisson Boltzmann Surface Area [49–51] (MM/PBSA), we further confirmed that ROCs, when bounded by GTP, were indeed less dimerized and therefore more prone to form the catalytically active monomers with hydrolysis activity than to be bound by its GDP counterparts.

Furthermore, we computed the mean first passage time (MFPT) for transition between the dominant conformations  $S3_{GDP}$  and  $S2_{GTP}$  based on the established MSM model (Figure 2F). According to the results, while  $S3_{GDP}$  can directly transit to  $S2_{GTP}$  relatively fast (143.39 ns), the time for transition from  $S2_{GTP}$  to  $S3_{GDP}$  was more than three times the reverse (446.58 ns). Additionally, compared with  $S3_{GDP}$ ,  $S2_{GTP}$  can more quickly transit to other states, which could be in line with the transient appearance of the active state during GTPase cycle. The unmatched interconversion rates between ROCs-GDP ( $S3_{GDP}$ ) and ROCs-GTP ( $S2_{GTP}$ ) observed in our simulation study, as exemplified by the considerably longer time taken from GTP-bound state back to GDP according to MFPT results, hinted that there could be undiscovered GTPase-activating proteins (GAPs) that help with accelerating this transition process under physiological conditions to keep this activation cycle in control.

Taken together, based on the ROCs-GDP~ROCs-GTP MSMs, although there is evidence claiming that ROCs form symmetric homodimers spontaneously, it appears that this process may not be a completely segregated from of nucleotide turnover, as GTP binding indeed induced prominent global conformational rearrangements in key structural motifs including Switch I, G4 loop, G5 loop, etc., while also rendering the whole ROCs complex less dimerized through altering the inter-monomer recognition pattern. In the meantime, the considerably longer time needed for transition from a GTP- to GDP-bound state also suggested the potential presence of GAPs as a catalyst for the ROC activation cycle [14,52].

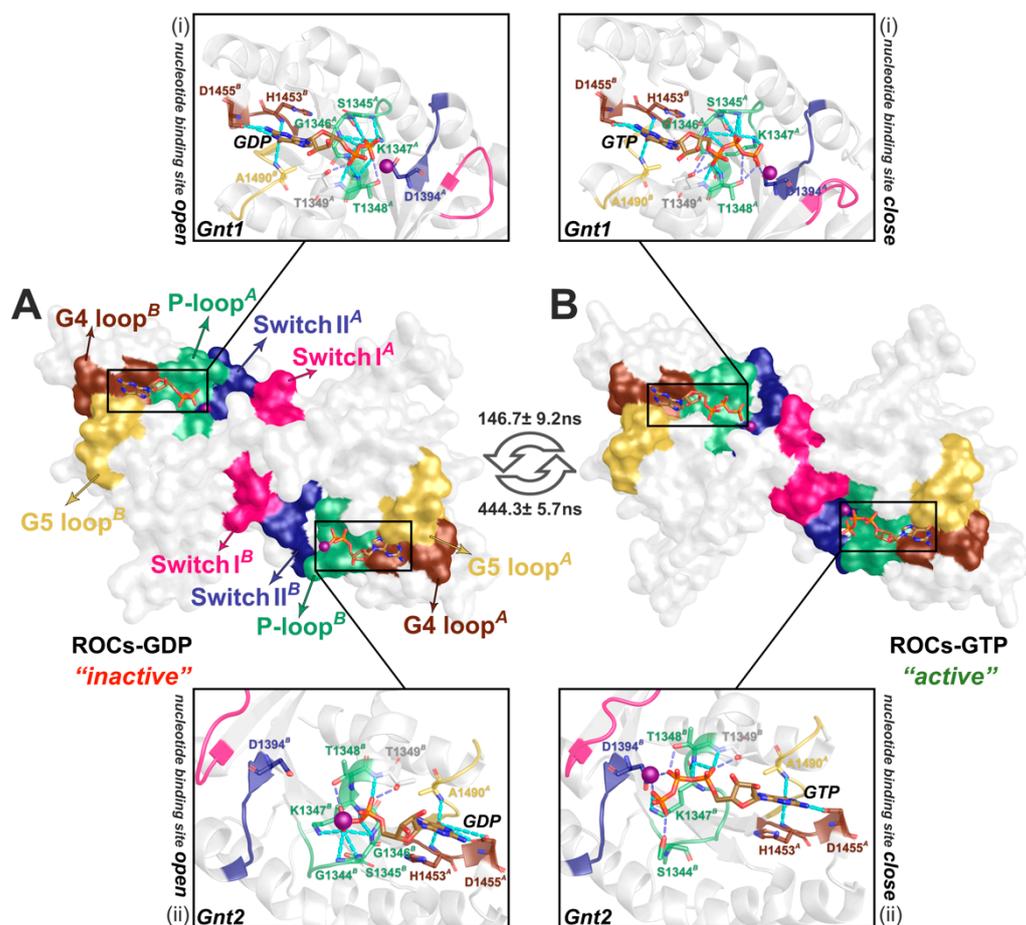
### 2.3. ROCs Exhibited Classic “Open” to “Closed” Conformational Transition of GTPase

As the primary function executed by ROCs, the GTPase activity is of interest due to their potential to fine-tune the bioactivity of the whole LRRK2 complex. Derived from the most-representative Ras from the small GTPase superfamily, whose nucleotide turnover typically leads to conformational transitions that is closely coupled with activation state, we also wonder if there would be considerable structural changes in regions contacting guanine nucleotide in the presence or absence of the  $\gamma$ -phosphate. With the identified representative metastable states in ROCs-GDP~ROCs-GTP MSM,  $S3_{GDP}$ ,  $S2_{GTP}$  represented the “inactive” and “active” state of ROCs (Figure 3). As previous studies suggested [20], the nucleotide-binding pockets on ROCs are contributed by regions from both monomers, and P-loop, Switch I, and Switch II, together with G4 and G5 loop from the other monomer, have been shown to play a key role (Figure 3). The results implied that the symmetric homodimerization pattern of ROC monomers rendered its two nucleotide-binding pockets indistinguishable in their interactions with GDP/GTP. For *Gnt1*, the guanine head in both GDP and GTP bonded with neighboring H1453<sup>B</sup> and D1455<sup>B</sup> of G4 loop, and also A1490<sup>B</sup> of G5 loop, while the polar residues S1345<sup>A</sup>-T1349<sup>A</sup> of P-loop interacted with the  $\alpha$ - and  $\beta$ -phosphate group of the nucleotide. Of note, this interaction network was stronger in the presence of GTP. Upon elongation of the phosphate tail with a  $\gamma$ -phosphate group by substituting GDP with GTP, T1348<sup>A</sup> formed additional interactions with the  $\gamma$ -phosphate, and D1394<sup>A</sup> of Switch II also hydrogen-bonded with the  $\gamma$ -phosphate; on the basis of enhanced interaction between GTP and neighboring residues, these together stabilized GTP binding and created a relatively “closed” conformation amenable for GTP hydrolysis (Figure 3B). Similar structural transitions were also observed in motifs mediating *Gnt2* binding. Such changes in regions implicated in nucleotide binding toward a “closed” conformation during the exchange of GDP with GTP in the activation cycle were also observed for classical GTPase including Ras [38], and their similarities could provide guidance on further investigations regarding the detailed GTP hydrolysis process by the ROC homodimer of LRRK2.

### 2.4. Nucleotide Turnover Reshaped ROCs Global Structure through Correlated Network

Generalized cross-correlation matrix [53] (GCCM), which is based on the fundamental definition of independence of random variables, was a well-established method to quantify both linear and nonlinear correlations and was thus employed to probe into the correlated motions that could be essential for ROCs functionality during the nucleotide shift. GCCM revealed that GTP substitution altered the correlative motion within ROCs (Figure 4A), including both inter- and intra-monomeric correlations. Generally speaking, the overall correlation network was significantly weakened in the presence of GTP, and this net effect was further demonstrated by reduced coupled dynamics between ROC<sup>A</sup> and ROC<sup>B</sup>, as well as within each monomer (highlighted in rectangles in Figure 4A). It is worth noting that inter-monomeric correlations dropped drastically, with the body subdomain, the primary part to provide an anchor for interactions stabilizing ROC dimerization, of each monomer exhibiting considerably weaker collective motion with the other monomer. Experiments have proposed that subdomains from ROC<sup>A</sup> and ROC<sup>B</sup> contributed to the formation of functional units ROCs1 and ROCs2 with a guanine nucleotide-binding site facilitated by polar interactions mediated by residues from both monomers. This decrease

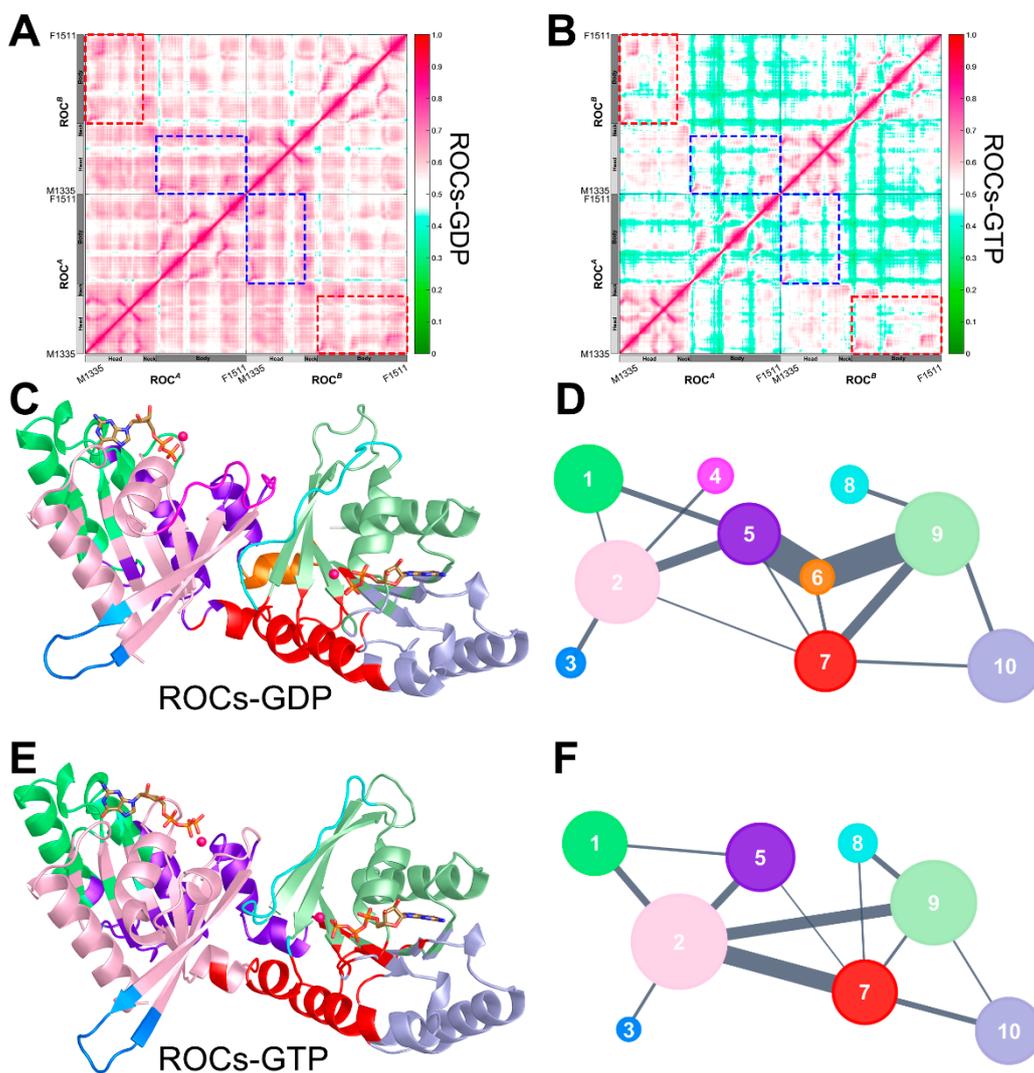
in inter-monomeric correlations indicated to a certain extent that the interactions between monomers that could be relevant to ROC dimerization were weakened upon substituting GDP with GTP. The observed change in GCCM correlation that mediated communication between distant regions was in line with the “open” to “closed” conformational shift during nucleotide exchange, and the weakened coupled motions between ROC monomers could be associated with the nucleotide-dependent dimerization-based activation of ROCs GTPase.



**Figure 3.** Detailed bonding network between the guanine nucleotide and ROCs in the representative structure from (A) ROCs-GDP,  $S3_{GDP}$ , and (B) ROCs-GTP,  $S2_{GTP}$ . The molecular surface diagram in the middle highlighted regions implicated in nucleotide binding, as color-coded in Figure 1B. The zoom-in graphs above (i) and below (ii) depicted the detailed interactions between ROCs and nucleotide *Gnt1* and *Gnt2*, respectively.

To investigate how intramolecular signals associated with nucleotide turnover propagate throughout the ROC complex to induce the observed shift in conformational landscape, we further conducted community network analysis (CNA) to illuminate the dynamic signal transduction scheme within ROCs during nucleotide exchange [54,55]. With CNA, through assigning nodes and weighting edges based on GCCM, closely related residues were grouped into inter-connected “communities”. The obtained community substructures for both ROCs-GDP and ROCs-GTP were shown in Figure 4C–F, and it turned out that the overall network partition was altered significantly upon nucleotide shift. In the presence of GDP (Figure 4C,D), the community structure of ROCs was more fragmented, and *Community* #6, consisting predominantly of ROC<sup>B</sup> neck residues, served as a hub for transmitting signals between the relatively distant head (*Community* #9) and body (*Community* #5) portion of ROC<sup>A</sup>. Meanwhile, strong information flow was also observed between ROC<sup>B</sup> head (*Community* #2) and ROC<sup>A</sup> body (*Community* #5), as well as ROC<sup>A</sup> head (*Community* #9) and ROC<sup>B</sup> body (*Community* #7). Additionally, 7 out of 10 communities were found

to encompass residues from both monomers (Table S4), demonstrating the considerably strong reciprocal interactions between the two monomers, which we believe to be relevant to dimerization. Upon substitution to GTP (Figure 4E,F), however, the number of communities was reduced to 8, while weaker intercommunications were found between communities. The head (*Community #2*) and body (*Community #7*) subdomain of  $ROC^B$  exhibited stronger and more direct connection, and although *Communities #2, #5, and #7* incorporated the neck residues of monomers, the original extensive connection network mediated by neck residues as an intramolecular “hub” in the presence of GDP no longer persisted. In the meantime, only five out of eight communities were still composed of residues from both monomers (Table S4), indicating weakened inter-monomeric crosstalk, which might facilitate the de-dimerization process of ROCs upon GTP binding to facilitate its activation.



**Figure 4.** Generalized cross correlation matrix (GCCM) of (A) ROCs-GDP and (B) ROCs-GTP. The colormap used for the graph is shown on the right with no/weak correlations indicated in green and strong ones in magenta. The correlation between  $ROC^A$  head and  $ROC^B$  body, and that between  $ROC^A$  body and  $ROC^B$  head subdomain, is highlighted in red and blue dashed line rectangles, respectively. Cartoon representation of (C) ROCs-GDP and (E) ROCs-GTP color-coded based on their corresponding community structures. Simplified community network map with inter-community crosstalk in (D) ROCs-GDP and (F) ROCs-GTP. Each community in the original structure is reduced to a sphere whose area is proportional to the residue components. Inter-community crosstalk is visualized by sticks of width that are proportional to edge connectivity between communities.

### 2.5. Molecular Basis and Allostery Underlying Pathogenetic Effects of R1441C/G/H Mutations

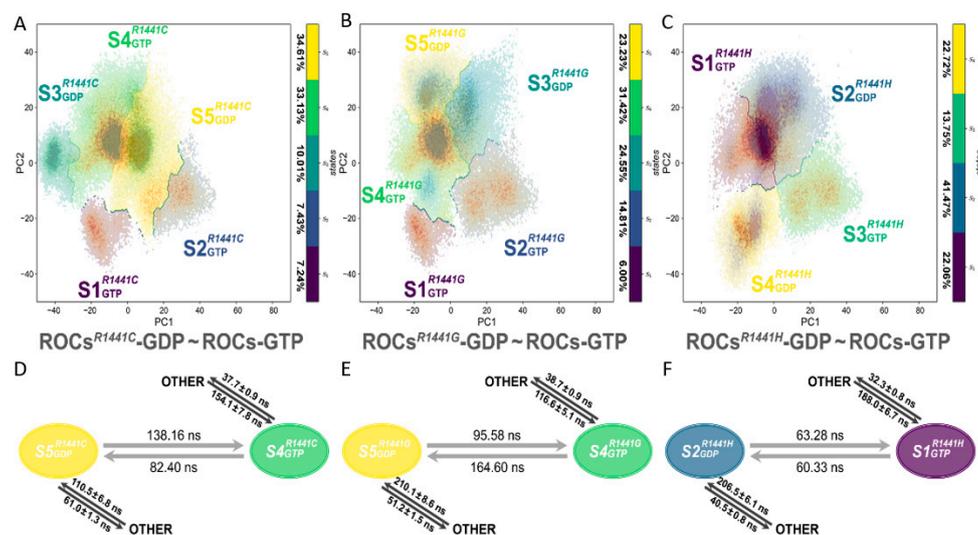
Belonging to the catalytic core of LRRK2, the ROC domain represents a hotspot for PD-relevant pathogenic mutations. Residue R1441, located at the dimerization interface of ROCs homodimer, engages in the exquisite interaction network to stabilize the complex. Moreover, mutations of R1441 (R1441C/G/H) have been proven to be disease-causing based on segregation with disease in PD families [30,56–58]. To understand the detailed mechanism accounting for pathogenic effects induced by R1441 point mutation, we mutated the arginine at position 1441 to cysteine, glycine, and histidine, respectively, in the starting structure of ROCs-GDP, and obtained the ROCs<sup>R1441C</sup>-GDP, ROCs<sup>R1441G</sup>-GDP, and ROCs<sup>R1441H</sup>-GDP models for further investigations. The three systems were also subjected to large-scale MD simulation as that for ROCs-GDP/ROCs-GTP, gathering 3  $\mu$ s extensive sampling for each mutated model. To probe into the disruption caused by R1441 mutations on the original conformational dynamics during nucleotide shift, we constructed kinetic MSM based on the data obtained with the same dimensionality-reduction transforming matrix and 3-ns lag time used for building ROCs-GDP~ROCs-GTP MSM, and yielded ROCs<sup>R1441C/G/H</sup>-GDP~ROCs-GTP MSMs with validated Markovianity, which could represent the conformational transition coupled with the activation of ROC homodimer under a mutated state (detailed in Section 4).

The whole conformational space was divided into 5, 5, and 4 metastables, respectively, for ROCs<sup>R1441C/G/H</sup>-GDP~ROCs-GTP MSMs (Figure 5A–C and Figure S4B–D) with a passing Chapman-Kolmogorov test (Figures S5–S7). In each MSM, based on the distribution pattern and fraction of the implicated systems in each metastable (Figure S4B–D, Table S1), we identified the most representative conformation for GDP-bound ROCs mutant and GTP-bound ROCs:  $S5_{GDP}^{R1441C}$  and  $S4_{GTP}^{R1441C}$  for the 5-metastable ROCs<sup>R1441C</sup>-GDP~ROCs-GTP MSM,  $S5_{GDP}^{R1441G}$  and  $S4_{GTP}^{R1441G}$  for the 5-metastable ROCs<sup>R1441G</sup>-GDP~ROCs-GTP MSM,  $S2_{GDP}^{R1441H}$  and  $S1_{GTP}^{R1441H}$  for the 4-metastable ROCs<sup>R1441H</sup>-GDP~ROCs-GTP MSM. The mutated ROCs<sup>R1441C/G/H</sup>-GDP shared much more overlap with ROCs-GTP on the PC1-PC2 plane (Figure 5A–C). MFPT was employed to estimate the relative conformational transition time between the GDP-bound ROCs mutant and GTP-bound ROCs (Figure 5D–F), and the results showed that mutation R1441C/G/H enabled a speeded shift from a GDP-bound inactive state to a GTP-bound active state (138.16, 95.58, and 63.28 ns compared with 143.39 ns) while also prominently accelerating the reverse transition (82.40, 164.60, and 60.33 ns compared with 446.58 ns). This observation supported the notion that R1441C/G/H mutation would result in overactivated ROCs, as this ROCs mutant could achieve faster conformational transition from a GDP-bound inactive state toward the GTP-bound active state. Moreover, inspection along the dimerization interface also disclosed the reduced interfacial area and disrupted bonding network mainly by the loss of the positively-charged R1441 sidechain upon mutation (Table S3). In the meantime, when measuring the dimerization binding free energy between monomers of these ROCs mutants with MM/PBSA (Table 1), a ~20 kcal/mol decrease was observed, demonstrating in energetic terms that ROCs harboring R1441C/G/H mutations were less dimerized and thus provided the basis for a faster transition toward the partially de-dimerized GTP-bound active state.

**Table 1.** ROC monomers MM/PBSA-binding free energy (kcal/mol) \*.

Systems	Representative	Binding Free Energy	Standard Deviations
ROCs-GDP	$S3_{GDP}$	−296.39	19.71
ROCs-GTP	$S2_{GTP}$	−285.09	23.52
ROCs <sup>R1441C</sup> -GDP	$S5_{GDP}^{R1441C}$	−276.31	18.02
ROCs <sup>R1441G</sup> -GDP	$S5_{GDP}^{R1441G}$	−274.93	17.48
ROCs <sup>R1441H</sup> -GDP	$S2_{GDP}^{R1441H}$	−274.98	17.64

\* The MM/PBSA-binding free energies between ROC monomers serve as an indicator for ROC dimerization extent.

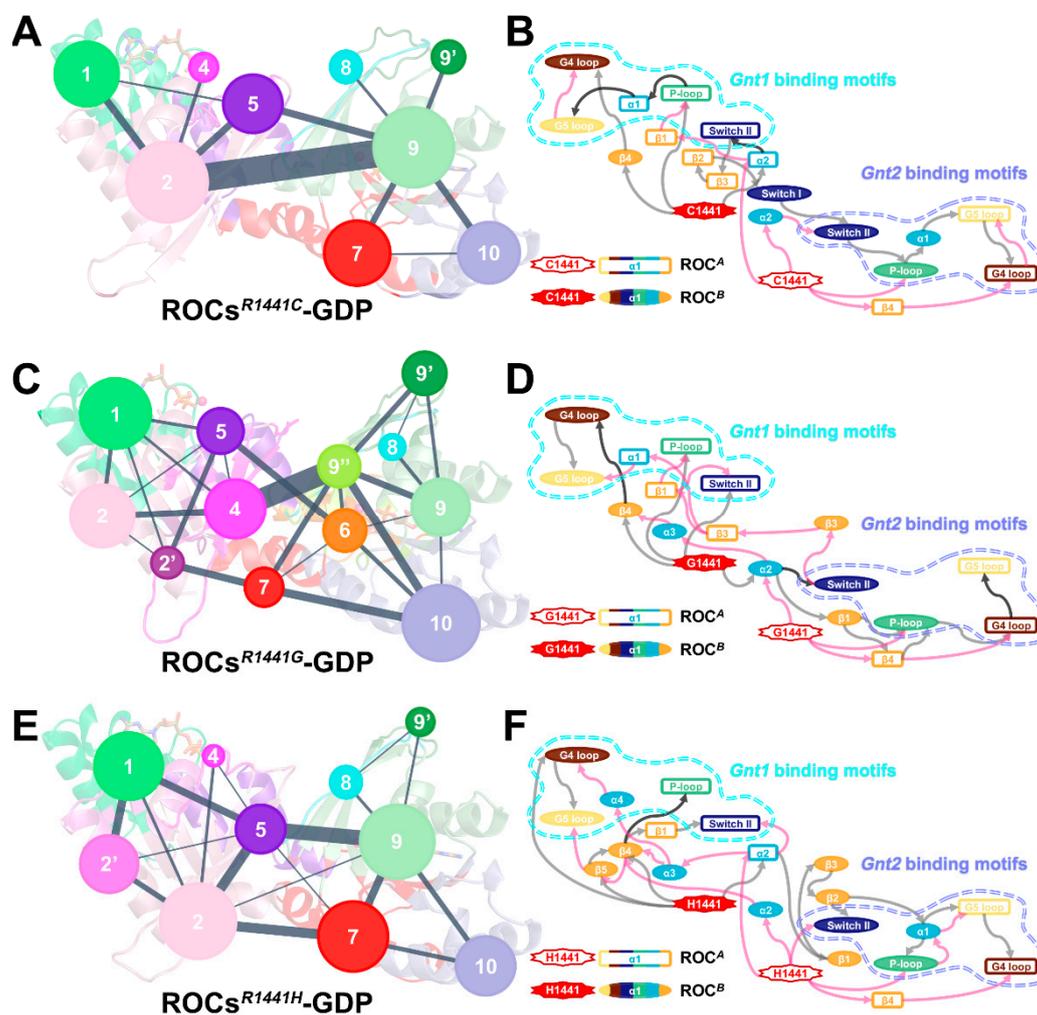


**Figure 5.** Distribution of metastable states in (A)  $\text{ROCs}^{R1441C}\text{-GDP}\sim\text{ROCs}\text{-GTP}$  (5 macrostates), (B)  $\text{ROCs}^{R1441G}\text{-GDP}\sim\text{ROCs}\text{-GTP}$  (5 macrostates), and (C)  $\text{ROCs}^{R1441H}\text{-GDP}\sim\text{ROCs}\text{-GTP}$  (4 macrostates) MSM. The type of nucleotide ligand for ROCs is indicated with subscripts, and the percentage of each macrostate is shown beside the corresponding colormap. MFPT between the predominant GDP-bound ROCs mutant and GTP-bound ROCs calculated in (D)  $\text{ROCs}^{R1441C}\text{-GDP}\sim\text{ROCs}\text{-GTP}$ , (E)  $\text{ROCs}^{R1441G}\text{-GDP}\sim\text{ROCs}\text{-GTP}$ , and (F)  $\text{ROCs}^{R1441H}\text{-GDP}\sim\text{ROCs}\text{-GTP}$  MSM. The two-ended black arrow between metastables and OTHER denoted the transition between the implicated macrostates and all other remaining metastable states in the corresponding MSM.

R1441 resides in  $\alpha 3$  of the body subdomain in both monomer, far from structural motifs mediating nucleotide binding. While the effects of R1441 mutations on ROC dimerization seemed relatively straightforward (through disrupting interaction network), how this point mutation altered the dynamics of distant regions, especially motifs associated with GDP/GTP binding, remained elusive. To address this problem, we employed GCCM and visualized the intramolecular correlation pattern in the mutated systems (Figure S8). We found that, compared with the wild-type ROCs dimer (Figure 4A), all mutants exhibited attenuated correlation, while the weakening of coupled dynamics mediated by the body subdomain where the mutation rests was particularly obvious. This proved that R1441C/G/H mutations altered the global conformational landscape through reshaping the intramolecular communication pattern. Together with the observed disruption of dimerization, it could thus be speculated that mutations at position R1441 might result in a conformational shift of the ROC's dimer toward certain states with a lower barrier to transmit into the GTP-bound active state [30].

The global impact induced by R1441C/G/H residue substitutions suggested that the mutations could act through an allosteric mechanism. Therefore, CNA was performed to gain valuable insights into the conformational reorganization and allosteric signaling upon mutation. Although the community structures of the three mutants have their own distinctive features (Figure 6A,C,E), one commonality was the uniformly weakened inter-community crosstalk, leaving relatively strong information flow only between communities from the non-body subdomains (*Community #2* and *Community #9* in  $\text{ROCs}^{R1441C}\text{-GDP}$ , *Community #4* and *Community #9* in  $\text{ROCs}^{R1441G}\text{-GDP}$ , and *Community #5* and *Community #9* in  $\text{ROCs}^{R1441H}\text{-GDP}$ ; community composition is detailed in Table S4). We subsequently analyzed the allosteric signaling path communicating mutation site and the nucleotide-binding regions with the *subopt* program from *Dynamic Network Analysis* [59–61] and visualized the results in terms of secondary structure (Figure 6B,D,F). Our results showed that the neck subdomain (mainly  $\alpha 2$ ) of both monomers was indispensable for transmitting signals induced by residue substitutions at R1441; meanwhile, this region was also implicated in propagating allosteric signals between monomers. Additionally,  $\beta 4$  from the body part of both monomers played a pivotal role in transmitting mutation-relevant allosteric signals

within each monomer. In wild-type systems (Figure S9), crosstalk between  $ROC^A$   $\alpha 2$  (neck) and  $ROC^B$   $\beta 3$  (head) bridged the functional units (ROCs1 and ROCs2) together in ROCs-GDP, while this signaling path was restructured upon GTP binding, with more segments participating in connecting the two functional units. For the mutants, the R1441C/G/H substitution also induced a change in signaling between functional units, as motifs from the body subdomain became directly involved in this process to enable a strengthened communication channel between distal regions, resembling the scheme observed in ROCs-GTP to some extent. In the meantime, the signaling network for ROCs<sup>R1441C/G/H</sup>-GDP is more concise with fewer roundabouts, indicating more efficient communication of allosteric signals from position 1441 to nucleotide-binding segments in mutants. Taken together, pathogenic mutations R1441C/G/H acted through an allosteric mechanism and resulted in a global conformational shift of ROCs' homodimer, which could enable a faster transition of the ROCs' mutant toward the active state.



**Figure 6.** Simplified community network map with inter-community crosstalk in (A) ROCs<sup>R1441C</sup>-GDP, (C) ROCs<sup>R1441G</sup>-GDP, and (E) ROCs<sup>R1441H</sup>-GDP. Cartoon diagrams with residues color-coded by their corresponding community partition were made transparent as the background to indicate the spatial position of each community. A schematic diagram of allosteric signaling between the mutated site at position 1441 and the nucleotide-binding motifs *Gnt1* and *Gnt2* (marked with cyan and orchid dotted lines) in (B) ROCs<sup>R1441C</sup>-GDP, (D) ROCs<sup>R1441G</sup>-GDP, and (F) ROCs<sup>R1441H</sup>-GDP mutants. Secondary structures of  $ROC^A$  are represented by colored hollow rectangles, and  $ROC^B$  components are shown in filled ovals. The gray and pink arrows, respectively, indicate allosteric signals from the mutation site on  $ROC^A$  and  $ROC^B$ , and black arrows denote paths involved in transmitting signals from both mutation sites. Note that all arrows here are one-ended for clarity; the actual information flow can pass in both directions for reversible communication.

### 3. Discussion

Mutations of LRRK2 are recognized as the most common genetic cause of PD. While a plethora of pathogenic mutations are found within the ROC domain, considerably less attention has been paid to understanding the importance of GTPase activity in regulating LRRK2 function, and we do not know anything regarding the regulation of the ROC GTPase cycle associated with activation [14]. While existing conventional experimental approaches failed to illuminate the dynamic machinery, large-scale MD simulations can provide invaluable insight into the originally unreachable dynamic process when integrating with MSMs. In the present study, based on MSMs, we aimed to unveil the nature of conformational transition of the ROCs' homodimer during GDP/GTP nucleotide turnover as well as quantitatively characterize the thermodynamic and kinetic properties. Through constructing ROCs-GDP~ROCs-GTP interconversion MSMs, we found that the ROCs' dimer exhibited a differed dimerization extent in the presence of GDP/GTP, which resembled the nucleotide-dependent oligomerization and activation of LRRK2 homologue *Chlorobium tepidum* Roco protein, and further supported the proposed regulation of ROC GTPase cycle through nucleotide-dependent dimerization. Interestingly, when estimating the relative transition time between inactive GDP-bound and active GTP-bound state, we found that the transition from ROCs-GDP to ROCs-GTP took only approximately 1/3 of the time needed for the reverse process, and this difference suggested the need for GTPase-activating proteins (GAP) that accelerate hydrolysis and render the complex inactive. Indeed, GAP-like proteins for LRRK2, as exemplified by ArfGAP1, have been reported [24,52,62]. The coupling between dimerization extent and nucleotide-binding state was further demonstrated by the observed conformational shift resembling the classic "open" to "closed" structural rearrangements observed for typical small GTPase like Ras [38], with regions implicated in nucleotide binding such as P-loop, Switch II, G4, and G5 loop, forming hydrogen bonds with the incoming  $\gamma$ -phosphate and thus yielding the "closed" conformation under a GTP-bound state. Probing into the correlation network that mediates the observed shift in conformational landscape, we found prominently attenuated inter-monomeric crosstalk upon GTP binding based on GCCM analysis, while a reshaped community structure with weakened signaling also suggested the potential oligomerization of ROCs-GTP during its GTPase cycle. On the whole, our present studies based on extensive MD simulations and MSMs suggested that the ROC domain of LRRK2 is likely to function through a nucleotide-dependent dimerization-based activation manner, similar to the GADs.

With our established activation model, we further investigated the mechanism underlying the pathogenic effects of missense mutations at R1441. A point mutation at this site was found to transform the intramolecular correlation pattern and community structure, resulting in faster conformational interconversion between the GDP-bound inactive and GTP-bound active state for a mutated ROCs homodimer. While previous studies have proved that R1441C/G/H mutations can both increase binding affinity for GTP and decrease GTPase activity to varying degrees, our present study expounded from the previously-unnoted aspect that such mutations were also likely to speed up the conformational transition during the GTPase activation cycle to trap ROCs in a more persistently active state, which would ultimately contribute to LRRK2 dysfunction and PD pathogenesis. In light of the effects of R1441C/G/H point mutations on the global conformational landscape and bioactivity of ROCs, we presumed that mutation-induced allostery played a key role in altering interdependent conformational dynamics within ROCs. Therefore, with allosteric signaling pathway analysis, structural motifs implicated in transmitting mutation-relevant allosteric signals were identified, which suggested the existence of potential allosteric pockets on the ROCs' homodimer for regulating its functionality. Taken together, the present study illuminated the dynamic conformational transition, which is coupled with nucleotide binding of the ROC domain of LRRK2, as a prototype for the ROCO proteins, during the GTPase activation cycle. Moreover, inspired by the disease-causing mutations within this domain, as exemplified by R1441C/G/H point mutations,

we revealed the underlying allostery of pathogenic effects, which would provide the basis for targeting such an allosteric mechanism to enable rational control of ROCs' bioactivity in both physiological and pathological contexts for PD treatment.

## 4. Materials and Methods

### 4.1. Construction of Simulation Systems

The five simulation systems (wildtype: ROCs-GDP and ROCs-GTP; mutant: ROCs<sup>R1441C</sup>-GDP, ROCs<sup>R1441G</sup>-GDP, and ROCs<sup>R1441H</sup>-GDP) in our study were built on the basis of the solved crystal structure of ROC homodimer in complex with GDP (PDB ID: 2ZEJ) [20]. The missing residues (T1357-D1387 and K1512-Q1516) in the original crystal profile were re-modeled with *Discovery Studio* [63], followed by subjecting the whole system to two rounds of 5000-step minimization using the steepest descent algorithm typed with a CHARMM forcefield [64]. The GTP-bound ROCs were obtained through introducing a  $\gamma$ -phosphate on the original GDP, and minimization was conducted on regions near the two nucleotide-binding sites to avoid potential structural conflicts. For the mutated systems, the arginine at position 1441 of both monomers in ROCs-GDP was mutated to cysteine (ROCs<sup>R1441C</sup>-GDP), glycine (ROCs<sup>R1441G</sup>-GDP), and histidine (ROCs<sup>R1441H</sup>-GDP), respectively.

### 4.2. MD Simulations

The parameter files for minimization and simulation were prepared with the Amber18 package using the Amber ff14SB force field and general amber force field (GAFF) [65,66]. All systems were first solvated in an orthorhombic transferable intermolecular potential three-point (TIP3P) water box [67], and Na<sup>+</sup> and Cl<sup>-</sup> counterions were subsequently added to attain the physiological saline concentration. Energy minimization was performed on all systems using the steepest descent and conjugate gradient algorithm [68,69] with protein scaffold fixed for 10 ps and without any constraint for 20 ps, respectively. All five systems were then put in a canonical ensemble (NVT) and were heated from 0 to 300 K within 300 ps with all protein atoms constrained, followed by being subjected to a 700 ps equilibration run. Finally, a total of 3  $\mu$ s (1  $\times$  3  $\mu$ s replica) conventional all-atom MD simulations was conducted on each system in an isothermal and isobaric ensemble (NPT) with periodic boundaries. Langevin dynamics using 1 ps<sup>-1</sup> collision frequency was chosen for the temperature control [70]. The Particle Mesh Ewald (PME) method was employed to analyze long-range electrostatic interactions [71], while a 10 Å cutoff was selected to treat short-range electrostatics and van der Waals forces. The SHAKE algorithm was used to constrain bond interactions involving hydrogens. The trajectory information of all systems was collected every 50 ps in each productive run.

### 4.3. Markov State Model Construction and Validation

Traditionally, to investigate the comprehensive conformational rearrangements in a dynamic context, large-scale MD simulations with few/no replications were often applied [72–74]. Such extensive sampling could allow access to slow dynamics, which often occur on a larger timescale and have been indicated to be responsible for intramolecular allosteric signaling [36,75]. However, due to the forbidding expense required to conduct such extensive simulations, the reproducibility of such work should be carefully assessed considering the lack of replications. In the meantime, integrating Markov state modeling (MSM) with MD simulations is gaining increasing popularity for the efficiency and accuracy that can be reached when interpreting biomolecular dynamics, and this combination has been proven reproducible when verifying with experimental techniques [22,39]. Nonetheless, it is clear that with very short trajectories, the intramolecular correlation network and also other important characteristics presented during the time-consuming conversion process would thus be uninvestigable. Under such circumstances, we constructed our statistical MSM using high-dimensional time-series data from several extensive simulations (3  $\times$  1  $\mu$ s), and this innovative scheme could allow the combination of the advantages of the aforementioned methodologies, achieving both the detailed atomistic insights of

traditional extensive simulations and statistical accuracy by integrating MSM only with short simulations.

We employed the Python library *PyEMMA* for the estimation, validation, and analysis of Markov state models (MSM) from MD data [42]. To understand the kinetics and thermodynamics of ROCs during nucleotide turnover under physiological conditions, we first built an MSM based on the simulation data from ROCs-GDP and ROCs-GTP systems (ROCs-GDP~ROCs-GTP MSM) to probe into their interconversion. First, the raw Cartesian data were featurized with inter-residue (represented by C $\alpha$  atoms) distances to describe the overall ROCs topology [45]. Subsequently, with the aim of capturing slow dynamics and primary conformational changes, we reduced the complexity of our data by carrying out dimensionality reduction with principal component analysis embedded in the *scikit-learn* python package [43,45,76,77]. Of note, we recorded the transforming matrix here for further application. Through implied timescales analysis, we validated the Markovianity of our data when discretized with 100 microstate cluster centers (Figure S2). After careful selection, a lag time of 3 ns and the maximum 200 *k-means* iteration number with 100 centers were chosen to cluster the data into microstates. The obtained microstates were then lumped into four metastable states using the PCCA+ algorithm with a passing Chapman-Kolmogorov test [78] (Figure S3). Transition pathway theory (TPT) was employed to measure the transition probability matrix of the constructed MSM [47,48,79] as well as compute the mean first passage time (MFPT) between selected states. The MSM describing kinetic interconversion between mutated ROCs<sup>R1441C/G/H</sup>-GDP and ROCs-GTP were also constructed in the same workflow with the recorded transforming matrix used for ROCs-GDP ~ROCs-GTP dimensionality reduction also employed here for the same purpose. Implied timescales and a Chapman-Kolmogorov test were also used to validate the Markovianity of the constructed models (Figures S2 and S5–S7).

#### 4.4. Generalized Cross-Correlation Matrix Analysis

Generalized cross-correlation matrix (GCCM) analysis, proposed by Grubmüller and Lange [53], was employed in our study for the assessment of both linear and nonlinear correlated motions. GCCM adopted the fundamental definition of independence of random variables, and treated variables  $x_i$ ,  $x_j$  correlated only when the product of their marginal distribution  $p(x_i) \cdot p(x_j)$  is larger than their joint distribution  $p(x_i, x_j)$ . Thus, mutual information (MI) between  $x_i$  and  $x_j$  is defined as:

$$MI[x_i, x_j] = \iint p(x_i, x_j) \ln \frac{p(x_i, x_j)}{p(x_i) \cdot p(x_j)} dx_i dx_j \quad (1)$$

which measures the degree of correlation between selected variables. Notably, the right part of Equation (1) is closely related to the well-known Shannon entropy  $H[x]$  [80], which can be obtained through:

$$H[x] = \int p(x) \ln p(x) dx \quad (2)$$

Therefore, the correlation between pairs of atoms  $x_i$  and  $x_j$  can be described with MI and calculated through the marginal Shannon entropy  $H[x_i]$ ,  $H[x_j]$ , together with the joint entropy term  $H[x_i, x_j]$  through:

$$MI[x_i, x_j] = H[x_i] + H[x_j] - H[x_i, x_j] \quad (3)$$

where the entropy terms  $H[x_i]$ ,  $H[x_j]$ , and  $H[x_i, x_j]$  can be calculated with the *g\_correlation* tool in Gromacs 3.3 [81]. The final generalized correlation coefficients ( $GC_{ij}$ ) can be obtained through normalizing  $MI[x_i, x_j]$  with:

$$GC_{ij} = \left\{ 1 - e^{-\frac{2MI[x_i, x_j]}{d}} \right\}^{\frac{1}{2}} \quad (4)$$

In which  $d$  represents the dimensionality of  $x_i$  and  $x_j$ . In our studies,  $d = 3$ .

#### 4.5. Dynamic Network Analysis

Concepts from network theories [41,82] were employed to investigate the residue–residue interaction network [41]. The whole ROCs homodimer of each system (except nucleotide and  $Mg^{2+}$  ion) was treated as a set of nodes (assigned to the C $\alpha$  atom of each residue) connected through edges, which would be drawn between nodes that stayed within a proximity of 4.5 Å for at least 75% of the simulation time. To indicate the relative distance for information transfer through a certain edge between nodes  $i$  and  $j$ ,  $GC_{ij}$  coefficients were used to weight the edges through:

$$d_{ij} = -\log GC_{ij} \quad (5)$$

The obtained  $GC_{ij}$ -weighted correlation network provides a basis for further community network analysis [54] (CNA) and allosteric pathway analysis [59–61]. The community substructure was obtained with the *gncommunities* program, which embedded the Girvan–Newman divisive algorithm and used “edge betweenness” that defines the number of shortest pathways across a given edge as an important partition criterion. To obtain the optimal substructure of the network, edges with the highest betweenness would be iteratively removed from the network, and the remaining ones would be recalculated until each node represents an isolated community.

The Floyd–Warshall algorithm [83] was used to identify the “shortest pathway” between selected pairs of nodes (defined as source and sink) by comparing the sum of  $d_{ij}$  of all edges implicated in the paths. In our studies, we treat R1441/C1441/G1441/H1441 as the source, while sinks were laid in regions associated with nucleotide binding, with G1346 representing P-loop, D1394 representing Switch II, D1455 representing G4 loop, and A1490 representing G5 loop.

**Supplementary Materials:** The following are available online, Figure S1: Real-time root-mean-square deviations (RMSD) of all systems during simulation, Figure S2: Implied timescales estimation of all constructed MSMs, Figure S3: Results of Chapman-Kolmogorov test when assuming 4 metastable states in ROCs-GDP~ROCs-GTP MSM, Figure S4: Distribution of the metastable states in all constructed MSMs, Figure S5: Results of Chapman-Kolmogorov test when assuming 5 metastable states in ROCs-GDPR1441C~ROCs-GTP MSM, Figure S6: Results of Chapman-Kolmogorov test when assuming 5 metastable states in ROCs-GDPR1441G~ROCs-GTP MSM, Figure S7: Results of Chapman-Kolmogorov test when assuming 5 metastable states in ROCs-GDPR1441H~ROCs-GTP MSM, Figure S8: GCCM matrix for mutated systems, Figure S9: Allosteric signaling path between R1441 and nucleotide binding regions in wildtype systems, Table S1: Percentage of each metastable structure in each MSM, Table S2: Bonding network formed between ROC monomers in ROCs-GDP and ROCs-GTP, Table S3: Interfacing area (Å<sup>2</sup>) and number of stable bonds for representative structures in each system, Table S4: Residue compositions of the identified communities in each system.

**Author Contributions:** Conceptualization, S.L., X.L., L.C. and X.C.; methodology, X.L. and Z.Q.; validation, X.L., Z.Q. and D.N.; formal analysis, X.L. and Z.Q.; investigation, X.L. and Z.Q.; resources, D.N.; data curation, X.L. and Z.Q.; writing—original draft preparation, X.L.; writing—review and editing, S.L.; visualization, S.L., X.L., L.C. and X.C.; supervision, S.L., L.C. and X.C.; project administration, S.L., L.C. and X.C.; funding acquisition, L.C. and X.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01), ZJ Lab, Shanghai Center for Brain Science and Brain-Inspired Technology, and the Shandong Provincial Natural Science Foundation (No. ZR201702200543).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Samples of the simulation trajectories are available from the authors upon request.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Samples of the intermediate structures mentioned in this work are available from the authors upon request.

## References

1. Wu, C.X.; Liao, J.; Park, Y.; Reed, X.; Engel, V.A.; Hoang, N.C.; Takagi, Y.; Johnson, S.M.; Wang, M.; Federici, M.; et al. Parkinson's disease-associated mutations in the GTPase domain of LRRK2 impair its nucleotide-dependent conformational dynamics. *J. Biol. Chem.* **2019**, *294*, 5907–5913. [[CrossRef](#)]
2. Petsko, G.A. The next epidemic. *Genome Biol.* **2006**, *7*, 108. [[CrossRef](#)] [[PubMed](#)]
3. Seol, W. Biochemical and molecular features of LRRK2 and its pathophysiological roles in Parkinson's disease. *BMB Rep.* **2010**, *43*, 233–244. [[CrossRef](#)] [[PubMed](#)]
4. Kalia, L.V.; Lang, A.E. Parkinson's disease. *Lancet* **2015**, *386*, 896–912. [[CrossRef](#)]
5. Halli-Tierney, A.D.; Luker, J.; Carroll, D.G. Parkinson Disease. *Am. Fam. Physician* **2020**, *102*, 679–691. [[PubMed](#)]
6. Moore, D.J.; West, A.B.; Dawson, V.L.; Dawson, T.M. Molecular pathophysiology of Parkinson's disease. *Annu. Rev. Neurosci.* **2005**, *28*, 57–87. [[CrossRef](#)]
7. Spillantini, M.G.; Schmidt, M.L.; Lee, V.M.Y.; Trojanowski, J.Q.; Jakes, R.; Goedert, M.  $\alpha$ -Synuclein in Lewy bodies. *Nature* **1997**, *388*, 839–840. [[CrossRef](#)]
8. Spillantini, M.G.; Goedert, M. Neurodegeneration and the ordered assembly of  $\alpha$ -synuclein. *Cell Tissue Res.* **2018**, *373*, 137–148. [[CrossRef](#)]
9. Simon, D.K.; Tanner, C.M.; Brundin, P. Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology. *Clin. Geriatr. Med.* **2020**, *36*, 1–12. [[CrossRef](#)]
10. Muda, K.; Bertinetti, D.; Gesellchen, F.; Hermann, J.S.; Von Zweyendorf, F.; Geerlof, A.; Jacob, A.; Ueffing, M.; Gloeckner, C.J.; Herberg, F.W. Parkinson-related LRRK2 mutation R1441C/G/H impairs PKA phosphorylation of LRRK2 and disrupts its interaction with 14-3-3. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E34–E43. [[CrossRef](#)]
11. Li, X.; Ye, M.; Wang, Y.; Qiu, M.; Fu, T.; Zhang, J.; Zhou, B.; Lu, S. How Parkinson's disease-related mutations disrupt the dimerization of WD40 domain in LRRK2: A comparative molecular dynamics simulation study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 20421–20433. [[CrossRef](#)]
12. Gasser, T. Molecular pathogenesis of Parkinson disease: Insights from genetic studies. *Expert Rev. Mol. Med.* **2009**, *11*, e22. [[CrossRef](#)]
13. Gasser, T. Mendelian forms of Parkinson's disease. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **2009**, *1792*, 587–596. [[CrossRef](#)]
14. Nguyen, A.P.T.; Moore, D.J. Understanding the GTPase activity of LRRK2: Regulation, function, and neurotoxicity. In *Advances in Neurobiology*; Springer: Berlin, Germany, 2017; Volume 14, pp. 71–88.
15. Marín, I.; Egmond, W.N.; Haastert, P.J.M. The Roco protein family: A functional perspective. *FASEB J.* **2008**, *22*, 3103–3110. [[CrossRef](#)] [[PubMed](#)]
16. Bosgraaf, L.; Van Haastert, P.J.M. Roc, a Ras/GTPase domain in complex proteins. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2003**, *1643*, 5–10. [[CrossRef](#)]
17. Zhang, P.; Fan, Y.; Ru, H.; Wang, L.; Magupalli, V.G.; Taylor, S.S.; Alessi, D.R.; Wu, H. Crystal structure of the WD40 domain dimer of LRRK2. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1579–1584. [[CrossRef](#)] [[PubMed](#)]
18. Rideout, H.J.; Stefanis, L. The neurobiology of LRRK2 and its role in the pathogenesis of Parkinson's disease. *Neurochem. Res.* **2014**, *39*, 576–592. [[CrossRef](#)]
19. Schmidt, S.H.; Weng, J.H.; Aoto, P.C.; Boassa, D.; Mathea, S.; Silletti, S.; Hu, J.; Wallbott, M.; Komives, E.A.; Knapp, S.; et al. Conformation and dynamics of the kinase domain drive subcellular location and activation of LRRK2. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2100844118. [[CrossRef](#)] [[PubMed](#)]
20. Deng, J.; Lewis, P.A.; Greggio, E.; Sluch, E.; Beilina, A.; Cookson, M.R. Structure of the ROC domain from the Parkinson's disease-associated leucine-rich repeat kinase 2 reveals a dimeric GTPase. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1499–1504. [[CrossRef](#)]
21. Wang, Y.; Ji, D.; Lei, C.; Chen, Y.; Qiu, Y.; Li, X.; Li, M.; Ni, D.; Pu, J.; Zhang, J.; et al. Mechanistic insights into the effect of phosphorylation on Ras conformational dynamics and its interactions with cell signaling proteins. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1184–1199. [[CrossRef](#)] [[PubMed](#)]
22. Lu, S.; Ni, D.; Wang, C.; He, X.; Lin, H.; Wang, Z.; Zhang, J. Deactivation Pathway of Ras GTPase Underlies Conformational Substates as Targets for Drug Design. *ACS Catal.* **2019**, *9*, 7188–7196. [[CrossRef](#)]
23. Vatansever, S.; Eрман, B.; Gümüş, Z.H. Comparative effects of oncogenic mutations G12C, G12V, G13D, and Q61H on local conformations and dynamics of K-Ras. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1000–1011. [[CrossRef](#)] [[PubMed](#)]
24. Biossa, A.; Trancikova, A.; Civiero, L.; Glauser, L.; Bubacco, L.; Greggio, E.; Moore, D.J. GTPase activity regulates kinase activity and cellular phenotypes of parkinson's disease-associated LRRK2. *Hum. Mol. Genet.* **2013**, *22*, 1140–1156. [[CrossRef](#)]
25. Ito, G.; Okai, T.; Fujino, G.; Takeda, K.; Ichijo, H.; Katada, T.; Iwatsubo, T. GTP binding is essential to the protein kinase activity of LRRK2, a causative gene product for familial Parkinson's disease. *Biochemistry* **2007**, *46*, 1380–1388. [[CrossRef](#)] [[PubMed](#)]

26. West, A.B.; Moore, D.J.; Choi, C.; Andrabi, S.A.; Li, X.; Dikeman, D.; Biskup, S.; Zhang, Z.; Lim, K.L.; Dawson, V.L.; et al. Parkinson's disease-associated mutations in LRRK2 link enhanced GTP-binding and kinase activities to neuronal toxicity. *Hum. Mol. Genet.* **2007**, *16*, 223–232. [[CrossRef](#)] [[PubMed](#)]
27. Cornejo-Olivas, M.; Torres, L.; Velit-Salazar, M.R.; Inca-Martinez, M.; Mazzetti, P.; Cosentino, C.; Micheli, F.; Perandones, C.; Dieguez, E.; Raggio, V.; et al. Variable frequency of LRRK2 variants in the Latin American research consortium on the genetics of Parkinson's disease (LARGE-PD), a case of ancestry. *NPJ Parkinson's Dis.* **2017**, *3*, 19. [[CrossRef](#)]
28. Ross, O.A.; Whittle, A.J.; Cobb, S.A.; Hulihan, M.M.; Lincoln, S.J.; Toft, M.; Farrer, M.J.; Dickson, D.W. Lrrk2 R1441 substitution and progressive supranuclear palsy. *Neuropathol. Appl. Neurobiol.* **2006**, *32*, 23–25. [[CrossRef](#)]
29. Weng, Y.H.; Chen, C.Y.; Lin, K.J.; Chen, Y.L.; Yeh, T.H.; Hsiao, I.T.; Chen, I.J.; Lu, C.S.; Wang, H.L. (R1441C) LRRK2 induces the degeneration of SN dopaminergic neurons and alters the expression of genes regulating neuronal survival in a transgenic mouse model. *Exp. Neurol.* **2016**, *275*, 104–115. [[CrossRef](#)]
30. Liao, J.; Wu, C.X.; Burlak, C.; Zhang, S.; Sahm, H.; Wang, M.; Zhang, Z.Y.; Vogel, K.W.; Federici, M.; Riddle, S.M.; et al. Parkinson disease-associated mutation R1441H in LRRK2 prolongs the "active state" of its GTPase domain. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4055–4060. [[CrossRef](#)] [[PubMed](#)]
31. Deyaert, E.; Wauters, L.; Guaitoli, G.; Konijnenberg, A.; Leemans, M.; Terheyden, S.; Petrovic, A.; Gallardo, R.; Nederveen-Schippers, L.M.; Athanasopoulos, P.S.; et al. A homologue of the Parkinson's disease-associated protein LRRK2 undergoes a monomer-dimer transition during GTP turnover. *Nat. Commun.* **2017**, *8*, 1008. [[CrossRef](#)] [[PubMed](#)]
32. Gasper, R.; Meyer, S.; Gotthardt, K.; Sirajuddin, M.; Wittinghofer, A. It takes two to tango: Regulation of G proteins by dimerization. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 423–429. [[CrossRef](#)] [[PubMed](#)]
33. Nixon-Abell, J.; Berwick, D.C.; Harvey, K. L'RRK de Triomphe: A solution for LRRK2 GTPase activity? *Biochem. Soc. Trans.* **2016**, *44*, 1625–1634. [[CrossRef](#)]
34. Dror, R.O.; Arlow, D.H.; Maragakis, P.; Mildorf, T.J.; Pan, A.C.; Xu, H.; Borhani, D.W.; Shaw, D.E. Activation Mechanism of the  $\beta$ 2-Adrenergic Receptor. *Biophys. J.* **2012**, *102*, 239a. [[CrossRef](#)]
35. Ni, D.; Li, X.; He, X.; Zhang, H.; Zhang, J.; Lu, S. Drugging K-RasG12C through covalent inhibitors: Mission possible? *Pharmacol. Ther.* **2019**, *202*, 1–17. [[CrossRef](#)]
36. Pal, S.; Kumar, V.; Kundu, B.; Bhattacharya, D.; Preethy, N.; Reddy, M.P.; Talukdar, A. Ligand-based pharmacophore modeling, virtual screening and molecular docking studies for discovery of potential topoisomerase I inhibitors. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 291–310. [[CrossRef](#)] [[PubMed](#)]
37. Meng, Y.; Shukla, D.; Pande, V.S.; Roux, B. Transition path theory analysis of c-Src kinase activation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9193–9198. [[CrossRef](#)]
38. Long, C.; Chao, E.; Da, L.-T.; Yu, J. A viral T7 polymerase ratcheting along DNA with fidelity control. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 638–644. [[CrossRef](#)]
39. Lu, S.; He, X.; Yang, Z.; Chai, Z.; Zhou, S.; Wang, J.; Rehman, A.U.; Ni, D.; Pu, J.; Sun, J.; et al. Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design. *Nat. Commun.* **2021**, *12*, 4721. [[CrossRef](#)]
40. Kokkonen, P.; Slanska, M.; Dockalova, V.; Pinto, G.P.; Sánchez-Carnerero, E.M.; Damborsky, J.; Klán, P.; Prokop, Z.; Bednar, D. The impact of tunnel mutations on enzymatic catalysis depends on the tunnel-substrate complementary and the rate-limiting step. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 805–813. [[CrossRef](#)] [[PubMed](#)]
41. Sethi, A.; Eargle, J.; Black, A.A.; Luthey-Schulten, Z. Dynamical networks in tRNA: Protein complexes. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6620–6625. [[CrossRef](#)] [[PubMed](#)]
42. Scherer, M.K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11*, 5525–5542. [[CrossRef](#)]
43. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
44. Prinz, J.H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J.D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105. [[CrossRef](#)] [[PubMed](#)]
45. Wu, H.; Noé, F. Variational Approach for Learning Markov Processes from Time Series Data. *J. Nonlinear Sci.* **2020**, *30*, 23–66. [[CrossRef](#)]
46. Peppas, N.A.; Robinson, J.R. Bioadhesives for optimization of drug delivery. *J. Drug Target.* **1995**, *3*, 183–184. [[CrossRef](#)] [[PubMed](#)]
47. Röblitz, S.; Weber, M. Fuzzy spectral clustering by PCCA+: Application to Markov state models and data classification. *Adv. Data Anal. Classif.* **2013**, *7*, 147–179. [[CrossRef](#)]
48. Weinan, E.; Vanden-Eijnden, E. Towards a theory of transition paths. *J. Stat. Phys.* **2006**, *123*, 503–523.
49. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508. [[CrossRef](#)]
50. Yang, Y.; Zhang, S.; Zhou, Q.; Zhang, C.; Gao, Y.; Wang, H.; Li, Z.; Wu, D.; Wu, Y.; Huang, Y.Y.; et al. Discovery of highly selective and orally available benzimidazole-based phosphodiesterase 10 inhibitors with improved solubility and pharmacokinetic properties for treatment of pulmonary arterial hypertension. *Acta Pharm. Sin. B* **2020**, *10*, 2339–2347. [[CrossRef](#)]

51. Zeb, A.; Son, M.; Yoon, S.; Kim, J.H.; Park, S.J.; Lee, K.W. Computational simulations identified two candidate inhibitors of Cdk5/p25 to abrogate Tau-associated neurological disorders. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 579–590. [[CrossRef](#)]
52. Stafa, K.; Trancikova, A.; Webber, P.J.; Glauser, L.; West, A.B.; Moore, D.J. GTPase activity and neuronal toxicity of parkinson's disease-associated LRRK2 is regulated by ArfGAP1. *PLoS Genet.* **2012**, *8*, e1002526. [[CrossRef](#)]
53. Lange, O.F.; Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins Struct. Funct. Bioinform.* **2006**, *62*, 1053–1061. [[CrossRef](#)]
54. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [[CrossRef](#)]
55. Zhang, Q.; Chen, Y.; Ni, D.; Huang, Z.; Wei, J.; Feng, L.; Su, J.-C.; Wei, Y.; Ning, S.; Yang, X.; et al. Targeting a cryptic allosteric site of SIRT6 with small-molecule inhibitors that inhibit the migration of pancreatic cancer cells. *Acta Pharm. Sin. B* **2021**, in press. [[CrossRef](#)]
56. Tong, Y.; Pisani, A.; Martella, G.; Karouani, M.; Yamaguchi, H.; Pothos, E.N.; Shen, J. R1441C mutation in LRRK2 impairs dopaminergic neurotransmission in mice. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14622–14627. [[CrossRef](#)] [[PubMed](#)]
57. Huang, X.; Wu, C.; Park, Y.; Long, X.; Hoang, Q.Q.; Liao, J. The Parkinson's disease-associated mutation N1437H impairs conformational dynamics in the G domain of LRRK2. *FASEB J.* **2019**, *33*, 4814–4823. [[CrossRef](#)] [[PubMed](#)]
58. Gilsbach, B.K.; Ho, F.Y.; Vetter, I.R.; Van Haastert, P.J.M.; Wittinghofer, A.; Kortholt, A. Roco kinase structures give insights into the mechanism of Parkinson disease-related leucine-rich-repeat kinase 2 mutations. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10322–10327. [[CrossRef](#)]
59. Del Sol, A.; Fujihashi, H.; Amoros, D.; Nussinov, R. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.* **2006**, *2*, 2006-0019. [[CrossRef](#)]
60. Ghosh, A.; Vishveshwara, S. A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15711–15716. [[CrossRef](#)]
61. Liang, Z.; Zhu, Y.; Long, J.; Ye, F.; Hu, G. Both intra and inter-domain interactions define the intrinsic dynamics and allosteric mechanism in DNMT1s. *Comput. Struct. Biotechnol. J.* **2020**, *23*, 749–764. [[CrossRef](#)]
62. Xiong, Y.; Yuan, C.; Chen, R.; Dawson, T.M.; Dawson, V.L. ArfGAP1 is a GTPase activating protein for LRRK2: Reciprocal regulation of ArfGAP1 by LRRK2. *J. Neurosci.* **2012**, *32*, 3877–3886. [[CrossRef](#)]
63. Studio, D. *Dassault Systemes BIOVIA, Discovery Studio Modelling Environment, Release 4.5*; Accelrys Softw Inc.: San Diego, CA, USA, 2015; pp. 98–104.
64. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)]
65. Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham, T., III; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M.; et al. *Amber18*; University of California: San Francisco, CA, USA, 2018.
66. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)]
67. Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960. [[CrossRef](#)]
68. Lu, S.; Chen, Y.; Wei, J.; Zhao, M.; Ni, D.; He, X.; Zhang, J. Mechanism of allosteric activation of SIRT6 revealed by the action of rationally designed activators. *Acta Pharm. Sin. B* **2021**, *11*, 1355–1361. [[CrossRef](#)] [[PubMed](#)]
69. Wójcik-Augustyn, A.; Johansson, A.J.; Borowski, T. Mechanism of sulfate activation catalyzed by ATP sulfurylase-Magnesium inhibits the activity. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 770–784. [[CrossRef](#)] [[PubMed](#)]
70. Stoltz, G.; Trstanova, Z. Langevin dynamics with general kinetic energies. *Multiscale Model. Simul.* **2018**, *16*, 777–806. [[CrossRef](#)]
71. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593. [[CrossRef](#)]
72. Palermo, G.; Ricci, C.G.; Fernando, A.; Basak, R.; Jinek, M.; Rivalta, I.; Batista, V.S.; McCammon, J.A. Protospacer Adjacent Motif-Induced Allostery Activates CRISPR-Cas9. *J. Am. Chem. Soc.* **2017**, *139*, 16028–16031. [[CrossRef](#)] [[PubMed](#)]
73. Palermo, G.; Miao, Y.; Walker, R.C.; Jinek, M.; McCammon, J.A. CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 7260–7265. [[CrossRef](#)]
74. Abramyan, A.M.; Yano, H.; Xu, M.; Liu, L.; Naing, S.; Fant, A.D.; Shi, L. The Glu102 mutation disrupts higher-order oligomerization of the sigma 1 receptor. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 199–206. [[CrossRef](#)] [[PubMed](#)]
75. Zhao, Q.; Ye, Z.; Su, Y.; Ouyang, D. Predicting complexation performance between cyclodextrins and guest molecules by integrated machine learning and molecular modeling techniques. *Acta Pharm. Sin. B* **2019**, *9*, 1241–1252. [[CrossRef](#)] [[PubMed](#)]
76. Pérez-Hernández, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **2013**, *139*, 015102. [[CrossRef](#)] [[PubMed](#)]
77. Molgedey, L.; Schuster, H.G. Separation of a mixture of independent signals using time delayed correlations. *Phys. Rev. Lett.* **1994**, *72*, 3634–3637. [[CrossRef](#)] [[PubMed](#)]
78. Tabar, M.R.R. Introduction to Stochastic Processes. In *Understanding Complex Systems*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 9–18.
79. Metzner, P.; Schütte, C.; Vanden-Eijnden, E. Transition Path Theory for Markov Jump Processes. *Multiscale Model. Simul.* **2009**, *7*, 1192–1219. [[CrossRef](#)]

- 
80. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 2005.
  81. Lindahl, E.; Hess, B.; Van Der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* **2001**, *7*, 306–317. [[CrossRef](#)]
  82. Mian, K.H. Interactive Dynamic Network Analysis. *Dyn. Soc. Netw. Model. Anal. Work. Summ. Pap.* **1982**, 117–119.
  83. Hougardy, S. The Floyd-Warshall algorithm on graphs with negative cycles. *Inf. Process. Lett.* **2010**, *110*, 279–281. [[CrossRef](#)]