

Article

An Insight of RuBisCO Evolution through a Multilevel Approach

Vladimir Camel^{1,2,3,*}  and Gaston Zolla^{1,2,4}

- ¹ Grupo de Investigación en Mutaciones & Biotecnología Vegetal, Facultad de Agronomía, Universidad Nacional Agraria La Molina, Av. La Molina s/n., Lima 12175, Peru; gzolla@lamolina.edu.pe
 - ² Doctoral Program in Biological Sciences and Engineering, National Agrarian University La Molina, Av. La Molina s/n., Lima 12175, Peru
 - ³ Programa de Fisiología Vegetal y Cambio Climático, Asociación ANDINUS, Calle Miguel Grau 370, Sicaya, Huancayo, Junín 12500, Peru
 - ⁴ Grupo de Investigación de Fisiología Vegetal, Facultad de Ciencias, Universidad Nacional Agraria La Molina, Av. La Molina s/n., Lima 12175, Peru
- * Correspondence: vlad_camel@hotmail.com

Abstract: RuBisCO is the most abundant enzyme on earth; it regulates the organic carbon cycle in the biosphere. Studying its structural evolution will help to develop new strategies of genetic improvement in order to increase food production and mitigate CO₂ emissions. In the present work, we evaluate how the evolution of sequence and structure among isoforms I, II and III of RuBisCO defines their intrinsic flexibility and residue-residue interactions. To do this, we used a multilevel approach based on phylogenetic inferences, multiple sequence alignment, normal mode analysis, and molecular dynamics. Our results show that the three isoforms exhibit greater fluctuation in the loop between αB and βC , and also present a positive correlation with loop 6, an important region for enzymatic activity because it regulates RuBisCO conformational states. Likewise, an increase in the flexibility of the loop structure between αB and βC , as well as Lys330 (form II) and Lys322 (form III) of loop 6, is important to increase photosynthetic efficiency. Thus, the cross-correlation dynamics analysis showed changes in the direction of movement of the secondary structures in the three isoforms. Finally, key amino acid residues related to the flexibility of the RuBisCO structure were indicated, providing important information for its enzymatic engineering.

Keywords: Bio3D; structural dynamics; structural flexibility; cross-correlation dynamics



Citation: Camel, V.; Zolla, G. An Insight of RuBisCO Evolution through a Multilevel Approach. *Biomolecules* **2021**, *11*, 1761. <https://doi.org/10.3390/biom11121761>

Academic Editor: Eric Sorin

Received: 23 September 2021
Accepted: 23 October 2021
Published: 25 November 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/>).

1. Introduction

RuBisCO (ribulose-1,5-bisphosphate carboxylase oxygenase) is the most abundant enzyme in nature and plays essential functions in the entry of carbon into the biosphere and in photorespiration processes [1]. It is found in most autotrophic organisms such as bacteria, archaea and eukarya (algae, higher plants) [2]. Evolutionary studies in RuBisCO have allowed its classification into four isoforms (I, II, III and IV) [3,4]. Isoform I is the predominant enzyme in nature and is found in cyanobacteria, green algae and in higher and lower plants. It is a holoenzyme consisting of eight large (RbcL) and eight small (RbcS) subunits [5]. The isoform II enzyme is present in bacteria and is composed only of large-type subunit multimers [(L₂)_x], and appears to be less efficient in cleaving CO₂ and O₂ [4,6]. Isoform II has a distinct physiological role, and it is used primarily to allow the Calvin–Benson–Bassham pathway to balance the cell redox potential [7,8]. Isoform III is found in archaeas and consists of a toroid-shaped pentagonal decamer composed of L subunits [9]. In addition, the enzyme shows extreme thermostability with high carboxylase activity at high temperatures [10,11] and exceeds the RuBisCO activity of spinach by 20 times, but it is not efficient at room temperature [12]. Moreover, it is not affected by the presence of oxygen [9,13,14]. Isoform IV includes proteins similar to RuBisCO (RLP) but does not use CO₂ as the main source of carbon [15]. Despite the variability of the amino acid sequences within the different RuBisCO isoforms [5,16], the key residues of the active site, catalytic chemistry and activation processes are conserved, and this supports the concept that there is a conserved set of

residues that are critical for folding and maintaining the general structure of the enzyme [15,17]. However, it is possible that proteins within the same isoform may have different enzymatic and kinetic properties. For example, phylogenetic studies show that the sequence of RbcL from *Arabidopsis thaliana* (5IU0) is different from *Oryza sativa* (1WDD) despite exhibiting high structural similarity. On the other hand, the amino acid sequences of isoform III in *Methanococcoides burtonii* (5MAC) is closer to isoform II. Likewise, the RuBisCO of *Nostoc* sp. (6KKM) and *Synechococcus elongatus* (6SMH) is isoform I, but their sequences are similar to isoform III. Due to these differences, it is necessary to understand the relationship between the structure and function of the RuBisCO enzyme in order to understand the role of the residues directly involved in catalysis. Furthermore, Nishitani et al. [10] showed that mutations (SP5-V330T) in the RbcL 3A12^{WT} protein of *Thermococcus kodakarensis* increased the flexibility of the α -helix 6 and loop 6 regions, being important to increase the photosynthetic efficiency of the enzyme at room temperature. Likewise, the closure of the active site implies movements of loop 6 and flexible elements of the N-terminal domain of the adjacent subunit in the dimer [18]. Currently, the two states, open and closed, of the RuBisCO enzyme are quite well-defined structurally, but the details of the closing mechanism are still unknown [19]. Therefore, it is necessary to study the influence between the structure, the amino acid composition and the flexibility of the RuBisCO structures.

On the other hand, RuBisCO is a widely studied enzyme. Consequently, the PDB (Protein Data Bank) repository has several RbcL structures [20], which are useful to understand the evolution of the different RuBisCO isoforms. In this sense, the Bio3D [21,22] and Prody packages emerged as computational tools that help to better understand the relationship between the structure, dynamics and function of sets of evolutionarily related proteins [23,24].

Consequently, in the present work, we evaluate how the evolution of sequence and structure among isoforms I, II and III of RuBisCO defines their intrinsic flexibility and residue-residue interaction.

2. Materials and Methods

2.1. Classification of RuBisCO Isoforms

RuBisCO protein codes (1RLC, 4RUB, 4HHH, 1GK8, 4LF1 and 3A12) were used to search for homologous structures using BLAST [25]; then, the different RuBisCO structures determined by crystallography were downloaded from RCSB PDB [20] using the Bio3D package [21,22]. A sequence identity threshold of 70% was used according to Kalenkiewicz et al. [26] to isolate structures of isoforms I, II and III. In this way 64 crystalline structures of the RbcL subunit of RuBisCO were downloaded, but redundant structures with missing amino acid residues were removed. These criteria allowed us to select 46 unique from wild-type RbcL and mutants in proteobacteria and archaea (Table S1 and Figure S1). The alignment of these amino acid sequences was performed with the MUSCLE algorithm [27]. All conformations were structurally superimposed on each other by least-squares fitting of the Cartesian coordinates of C- α atoms equivalent to the C-terminal domain, since this region was found to be the most structurally invariant. Principal component analysis (PCA) was used to evaluate the relationships between conformer sets of overlapping structures, as it is very useful for evaluating the distributions of experimental structures and comparing them with the conformations obtained through molecular dynamics (MD) simulations (Figure S1) [21,26].

2.2. RuBisCO Structure Selection and Phylogenetic Analysis

Based on the component analysis, RbcL^{WT} structures and mutants from model organisms were selected which had a resolution ≤ 2.7 Å [28,29], which did not have missing amino acid residues, and for which the crystallized structure was $\geq 95\%$ of the total protein (RbcL) in its three isoforms (I, II, III). Thus, 137 RuBisCO protein sequences, including the sequences provided by Kacar et al. [7] and the selected RuBisCO sequences, were used to build a phylogenetic tree according to Kacar et al. [7] with the PhyloBot web service [30]. RuBisCO orthologs were identified by the NCBI BLAST tool [25]. Then, multiple sequence alignments

were inferred by the MSAProbs [31] and MUSCLE [27] algorithms with their default settings. Maximum likelihood (ML) phylogenetic inference was estimated using the PROTCATWAG model [32,33] in the RAxML web service [34]. Subsequently, ML phylogeny files were exported to the PhyML website [35] in order to calculate statistical support for branches as approximate likelihood ratios and the sequence from the group IV family as the outgroup to root the tree [7]. Finally, the phylogeny plot was developed with Mega6 software [36].

2.3. Normal Mode Analysis

Normal mode analysis (NMA) is a simple method to predict and characterize the internal dynamics of proteins, where slow low-frequency movements are often of functional importance [21,22]. NMA analyses were developed in the Bio3D package, where simultaneous analysis of a large set of structures is easily performed through the implementation of ensemble normal mode analysis (eNMA) [21,22], allowing the rapid characterization and comparison of flexibility across homologous structures. eNMA allows the prediction and identification of different flexibility patterns between different protein isoforms that are available at PDB [21,22]. In this way, high resolution crystallographic structures of the RbcL subunits of RuBisCO were selected: 6 structures of isoform I (PDB code: 1WDD, 4RUB, 5IU0, 1IWA, 1GK8 and 6FTL), 3 structures of isoform II (PDB code: 4LF1^{WT}, 5HAN^{S59F} and 5HJX^{A47V}) and 3 structures of isoform III (PDB code: 3A12^{WT}, 3KDO^{SP6} and 3WQP^{T289D}). As input, the set of pdb structures aligned with MUSCLE software was provided [27]. Then, an efficient model based on C-alpha was used to enable the modes to be calculated quickly. Aligned eigenvectors and mode fluctuations were obtained as results for all RbcL structures.

2.4. Molecular Dynamics

The simulation models were built based on the high-resolution crystallographic structures of the RbcL subunits of RuBisCO; 6 structures of the isoform I were selected (PDB code: 1WDD, 4RUB, 5IU0, 1IWA, 1GK8 and 6FTL), as well as 3 structures of isoform II (PDB code: 4LF1^{WT}, 5HAN^{S59F} and 5HJX^{A47V}) and 3 structures of isoform III (PDB code: 3A12^{WT}, 3KDO^{SP6} and 3WQP^{T289D}). Before running MD simulations, proteins were treated. First, water molecules and monomers were eliminated. Moreover, missing amino acid residues from all structures were completed with MODELLER software version 10.0 (Accelrys, San Diego, CA, USA) [37]. The selected models satisfied spatial constraints such as bond lengths, bond angles, dihedral angles, and interactions between unbound residues. Models' stereochemical quality was assessed with Ramachandran graphs generated on the MolProbity server [38], and fold quality was determined by Verify3D [39,40].

Molecular dynamics simulation was performed using the Groningen Machine for Chemical Simulations GROMACS version 2020 [41]. The PDB2GMX module was used to generate the topology that had information about the unbound parameters (types of atoms and charges) and bound parameters (bonds, angles and dihedrals) within the simulation. The CHARMM36 force field [42] was used for the simulations of all RuBisCO systems following similar studies [14,43,44]. Periodic boundary conditions (PBC) were applied in all directions of a cube box with a 10 Å lateral size. The systems were solvated with the TIP3P water model [36]. Na⁺ ions were added to neutralize the system, as in previous studies [45,46]. To minimize energy in all systems, the algorithm of descending steps was used with 50,000 steps and with a search for energy less than 1000 kcal/mol. We used the isothermal-isobaric set with two equilibrium phases to simulate a system at cellular physiological conditions. The first equilibrium phase was done in the NVT ensemble at a constant temperature of 300 K with a Berendsen thermostat. The second equilibrium phase was done in the NPT ensemble at a pressure of 1 bar for 2 ns with the Parrinello–Rahman barostat. The simulation was carried out for 50 ns with integration steps of 2 fs under constant pressure and temperature conditions with the leapfrog integration algorithm. The LINCS algorithm was used to constrain all bonds during equilibrium [47], and the Ewald particle mesh algorithm was used for long-range ionic interactions.

In the 50-ns MD simulation, 5000 trajectories were obtained. The analysis of the output structures was performed by the following GROMACS commands: `gmx_mpi rmsd` to calculate root mean square deviation (RMSD) values; `gmx_mpi rmsf` to calculate root mean square fluctuation (RMSF) values; and `gmx_mpi gyrate` to calculate the radius of gyration. Finally, PCA and DCCM analyses were carried out with the Bio3D package, following Yu and Dalby's [48] recommendations. Conversion of the trajectory from XTC to DCD format was done with the CatDCD plugins of VMD software [49], and Pymol was used for image editing [50].

2.5. Stability and Flexibility Analysis

RMSD was used to measure the deviations of the protein backbone from its original structural conformation to its final structural conformation. When the stationary phase of the RMSD curve is reached, the protein is in equilibrium [51]. On the other hand, RMSF was used to measure the average individual residue flexibility during MD simulation. RMSF can indicate structurally which amino acids in a protein are more important in molecular motion [51]. RMSD and RMSF were performed using built-in protocols from GROMACS [41] and Bio3D [21,22].

2.6. Principal Component Analysis

Principal component analysis (PCA) was performed using the Bio3D package [21] implemented in the R-Project and ProDy software [23,24]. The PCA was carried out on C α atoms during the last 40 ns of the trajectories [52]. The correlated movements of the whole protein can be represented by the eigenvectors and eigenvalues. The eigenvectors, also called principal components (PC), gave the direction of the coordinated movement of the atoms, and the eigenvalues represented the magnitude of the movement along the corresponding eigenvectors [53]. Thus, PC1 and PC2 were computed, because they contributed more significantly to the PCA analysis [54].

Briefly, the PCA was based on the diagonalization of the covariance matrix, C , with elements C_{ij} calculated from the aligned and overlapping Cartesian coordinates, r , of equivalent C α atoms [21]:

$$C_{ij} = \langle r_i - \langle r_i \rangle \cdot (r_j - \langle r_j \rangle) \rangle$$

where r_i and r_j are cartesian coordinates of the i th and j th C α atoms, and $\langle r_i \rangle$ and $\langle r_j \rangle$ represent the average time over all configurations derived from the molecular dynamics simulation. The analysis was limited to C α atoms because they were less disturbed by statistical noise and offers a meaningful characterization of essential spatial movements [55].

2.7. Dynamic Cross-Correlation Matrices (DCCM)

To have a better understanding of the dynamics of the three RuBisCO isoforms, cross-correlation analysis (DCCM) was used to evaluate the motions (shifts) of alpha (C α) carbon atoms in the MD simulations of all systems [56]. Additionally, it provides useful information regarding the mutation effect on protein dynamics by analyzing how atomic shifts were correlated [57,58], and it was constructed using the Bio3D package from R-Project [21].

The DCCM map is a 3D matrix annotation that displays time-related information for protein residues. Time-dependent data based on residuals can be analyzed using visual pattern recognition. The DCCM map shows the correlations of amino acid movements, and was calculated according to Ichiye and Karplus' [59] equation:

$$C_{ij} = \frac{(\Delta r_i \cdot \Delta r_j)}{(\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{1/2}}$$

where Δr_i and Δr_j are the displacements from the mean position of the i -th and j -th atoms with respect to time. The angle brackets " $\langle \rangle$ " represent the average time over the entire trajectory. C_{ij} values ranged from -1 to $+1$; a positive value represented a positively

correlated movement between residues i and j , while a negative value implied a negatively correlated movement between residues i and j [60,61].

3. Results

3.1. RuBisCO Forms Classification

From the RCSB protein database, 64 crystal structures of the RbcL subunit of RuBisCO were downloaded. A total of 18 structures were not considered due to a lack of coordinates, leaving 46 RuBisCO complexes between wild-types and mutants (Figure 1). Next, principal component analysis (PCA) was carried out. 73% of the total variance of the atomic fluctuations was captured along the first principal component (PC), while the second and third dimensions were necessary to capture 83.2 and 88.2, respectively (Figure 1).

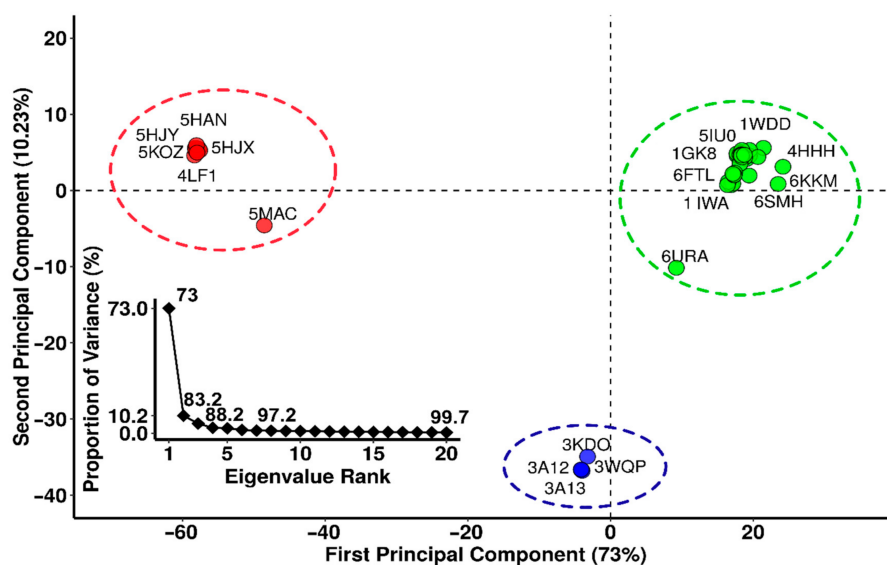


Figure 1. Principal component analysis (PCA) of RuBisCO RbcL isoforms. Green represents isoform I, red represents isoform II and isoform III is in blue.

The PCA structure shows three conformational clusters of RuBisCO (Figure 1). The largest cluster (in green) corresponds to 32 proteins (1BXN, 5OYA, 6FTL, 5MZZ, 5NV3, 1IWA, 1BWV, 5WSK, 2V6A, 1UW9, 2V68, 2VDH, 2VDI, 1GK8, 2V69, 4HHIH, 4RUB0, 4MKV, 1IR1, 1WDD, 3ZXW, 1RSC, 1RBL, 7JFO, 6URA, 1SVD, 1RLC, 1RLD, 1EJ7, 6SMH and 6KKM) and involves most RuBisCO structures from higher plants, green algae, blue-green algae, cyanobacteria, diatoms, and proteobacteria, including the 6URA structure of the bacteria *Promineofilum breve*, which is a benchmark to understand the evolution of RuBisCO Form I (Table S1). The second cluster (in red) includes 10 proteins (5MAC, 4LF1, 5HQM, 5KOZ, 5HQL, 5HJY, 5HAT, 5HAN, 5HAO and 5HJX), and they are mainly proteobacteria that present RuBisCO Form II, with the exception of 5MAC, which is found in *Methanococcoides burtonii* (archaea) and represents form II/III (Figure 1). Finally, in the third cluster (in blue), there are 4 proteins (3KDO, 3A13, 3WQP and 3A12), which corresponded to archaea (Figure 1).

Based on PCA analysis, RbcL wild-type structures and mutants from model organisms (Figure 1) were selected with a resolution ≤ 2.7 Å, with crystallized structure $\geq 95\%$ and without any lost amino acid residues. This criterion allowed 6 wild-type structures to be selected from the largest cluster (in green: 1WDD, 5IU0, 4RUB, 1GK8, 6FTL and 1IWA). From the second cluster (in red) 1 wild-type structure was selected (4LF1^{WT}) as well as 2 mutants (5HJX^{A47V} and 5HAN^{S59F}), and from the third cluster (in blue) 1 wild-type (3A12^{WT}) and 2 mutants (3KDO^{SP6} and 3WQP^{T289D}) were selected.

To classify the 12 selected structures according to their evolutionary groups, a maximum likelihood (ML) phylogenetic tree was made with 137 amino acid sequences of RuBisCO RbcL. The phylogeny shows that the different RuBisCO isoforms (I, II, III and

IV) share a common evolutionary ancestor. The RbcL sequences of isoform I have subgroups IA, IB, and IC/D, which includes higher plants, cyanobacteria, green algae, red algae, and some bacteria. Consequently, the RuBisCO structures of the *Arabidopsis thaliana* (5IU0), *Nicotiana tabacum* (4RUB), *Oriza sativa* (1WDD) and *Chlamydomonas reinhardtii* (1GK8) species corresponded to IB subgroup (Figure 2), while the *Galdieria partita* (1IWA) and *Skeletonema marinoi* (6FTL) are related to the IC/D subgroup (Figure 2). The RbcL sequences of isoform II include proteobacteria and some eukaryotic alveolates; also within this clade are *Rhodospseudomonas palustris*, including members of this species which are wild-types (4LF1^{WT}) and mutants (5HJX^{A47V} and 5HAN^{S59F}) (Figure 2). Other wild-type (3A12^{WT}) and mutant (3KDO^{SP6} and 3WQP^{T289D}) structures belong to *Thermococcus kodakarensis*; its lineage arches and its RuBisCO structure are form III (Figure 2).

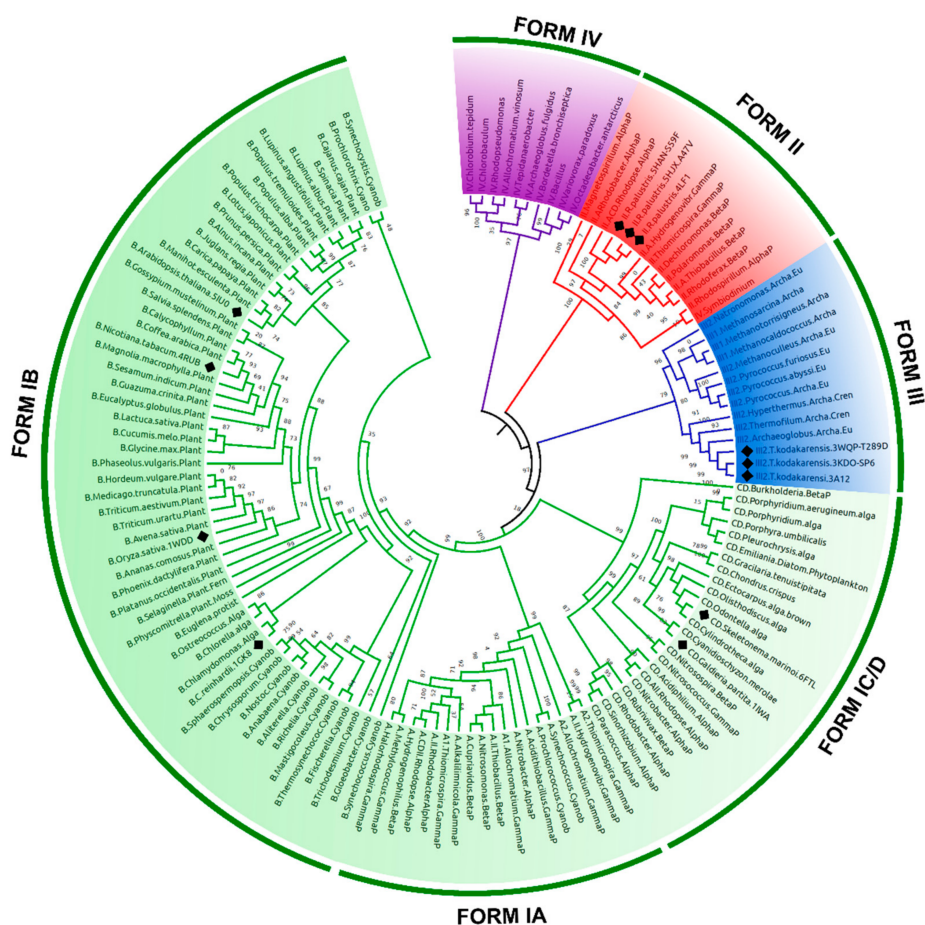


Figure 2. Maximum likelihood phylogenetic analysis of the RuBisCO RbcL protein family. Four forms of RuBisCO were classified along an evolutionary trajectory from the most recent common ancestor. Ancestral sequences were tagged according to their position in the RuBisCO subfamilies. Phylogenetic positions of the selected Wt and mutant proteins were marked with a black rhombus.

According to multiple sequence alignment analysis, the ends of the N-terminal and C-terminal regions showed greater variation in amino acid sequences. In Figure 3, the conserved active site residues are shown with an asterisk in the alignment. The secondary structures that maintain the amino acid residues involved in the catalysis were: α B (E72), α 0 (N144), a loop that connects β 1 and α 1 (K202; K204), a catalytic motif located between α 1 and α 2 (G223; D225; F226; K228; D230; E231), β 5 (H324) and loop 6 (K366) (Figure 3). Likewise, the secondary structures that are conserved and involved in the union of the phosphate groups in C1 and C5 of RuBP were: a loop between α B and β C (T77), β 5 (R325), β 6 (H358), β 6, a loop that connects β 7 and α 7 (S411; G413) and a loop connecting β 8 and α 8 (G436; G437) (Figure 3).

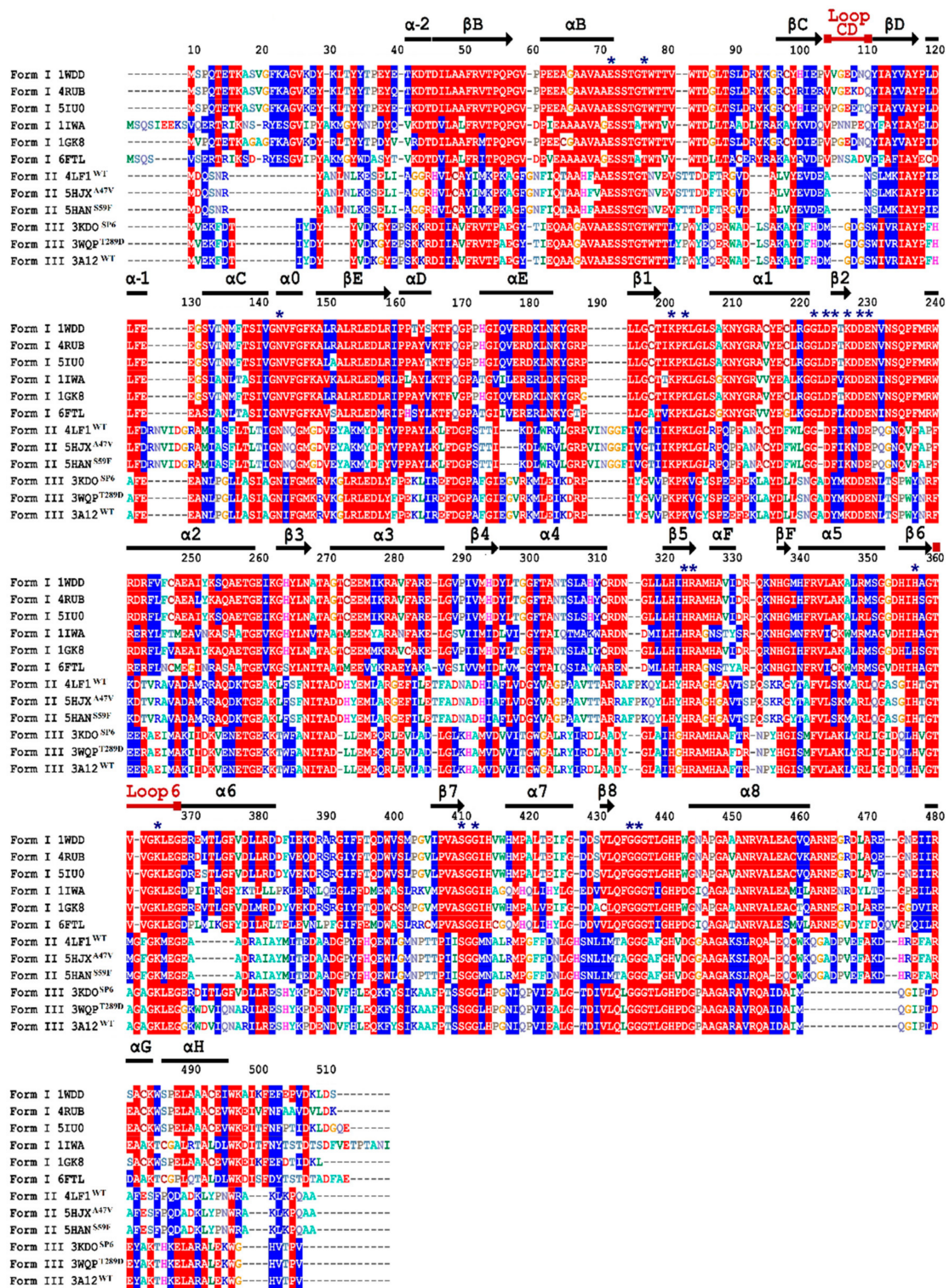


Figure 3. Multiple sequence alignment based on the structure of three isoforms of RuBisCO. Isoform I: *O. sativa* (1WDD), *N. tabacum* (4RUB), *A. thaliana* (5IU0), *C. reinhardtii* (1GK8), *G. partita* (1IWA) and *S. marinoi* (6FTL). Isoform II: 1 wild-type (4LF1^{WT}) and 2 mutants (5HJX^{A47V} and 5HAN^{S59F}) of *R. palustris*. Isoform III: 1 wild-type (3A12^{WT}) and 2 mutants (3KDO^{SP6} and 3WQP^{T289D}) of *T. kodakarensis*. Amino acid residue numbers are shown at the top of the sequence. Secondary structural elements, such as α -helices (bars) and β -strands (arrows), are shown in the figure. With an asterisk is depicted, it indicates mechanically important active site residues. Red background shading represents identical amino acids, blue shading designates similar amino acids while white shading indicates no similarity.

The catalytic loop 6 in Figure 3 is characterized as a conserved and flexible sequence because it interacts with the tail in the C-terminal region to close on the catalytic pocket when it binds to RuBP. This tail then opens to allow product release. On the other hand, the CD loop is located in the N-terminal domain and approaches the opening of the active site from the opposite direction to loop 6; furthermore, it is packed against loop 6. The observed differences among RbcL sequences and between species is reflected in the molecular complexity of RbcL isoforms (Figure 3).

3.2. Stability and Flexibility Evaluation of RuBisCO Forms

Protein function depends on its structure and dynamics and can be altered by mutations. Consequently, it is necessary to understand the intrinsic structural flexibility of the observed differences in multiple sequence alignment. Thus, the flexibility of the RuBisCO forms was evaluated by normal mode analysis (NMA) and molecular dynamics (MD). In Figure 4, the eNMA shows the consensus fluctuations are highlighted and reveal a conserved pattern among species and RuBisCO forms (Figure 4a,b). The three isoforms show a greater fluctuation in the N-terminal domain spanning the amino acid residues (51–68) between the secondary elements α B and β C (Figure 4a), which are functionally relevant for RbcL. Likewise, RuBisCO form III presents greater fluctuation (≥ 3 Å) with respect to form I and II (Figure 4a,b). Conversely, the catalytic domain of the α/β barrel subunit (150–444) was more stable in all structures evaluated.

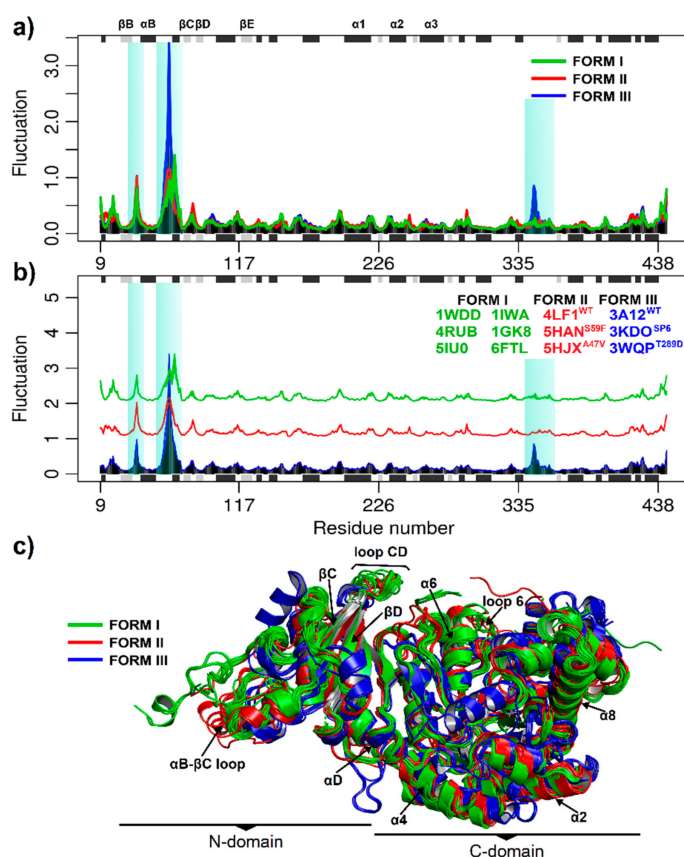


Figure 4. Normal mode analysis. (a) Consensus fluctuations of RuBisCO forms I, II and III. α -helices are in black and β -strands in gray; (b) Consensus fluctuations among species. In Form I are *O. sativa* (1WDD), *N. tabacum* (4RUB), *A. thaliana* (5IU0), *C. reinhardtii* (1GK8), *G. partita* (1IWA) and *S. marinoi* (6FTL). In form II are *R. palustris* with 1 wild-type (4LF1^{WT}) and 2 mutants (5HJX^{A47V} and 5HAN^{S59F}), and in form III are *T. kodakarensis* with 1 wild-type (3A12^{WT}) and 2 mutants (3KDO^{SP6} and 3WQP^{T289D}); (c) Monomeric structure of RbcL. The monomer is divided into two domains: the N-terminal domain and the C-terminal domain. Loop α B- β C, loop CD, loop 6 and α 6 are indicated. The different colors indicate Form I (green), II (red), and III (blue).

To evaluate the conformational changes of RuBisCO isoforms, we performed MD analysis for 50 ns over time. The mean square deviation (RMSD) was used to evaluate the conformational stability of the protein during the simulations. The mean square fluctuation (RMSF) was useful to identify rigidity and flexibility among RuBisCO forms. An RMSF value greater than 0.3 nm was considered as high fluctuation [62]. In Figure 5a, RMSD showed great variability in form I, where 4RUB (~0.36 nm ± 0.00268), 1GK8 (~0.34 nm ± 0.0025), 6FTL (~0.37 nm ± 0.0027) and 1IWA (~0.36 nm ± 0.0032) presented the lowest values (Table 1). Contrary to this, 1WDD (~0.52 nm ± 0.004) and 5IU0 (~0.46 nm ± 0.0033) showed very high RMSD values.

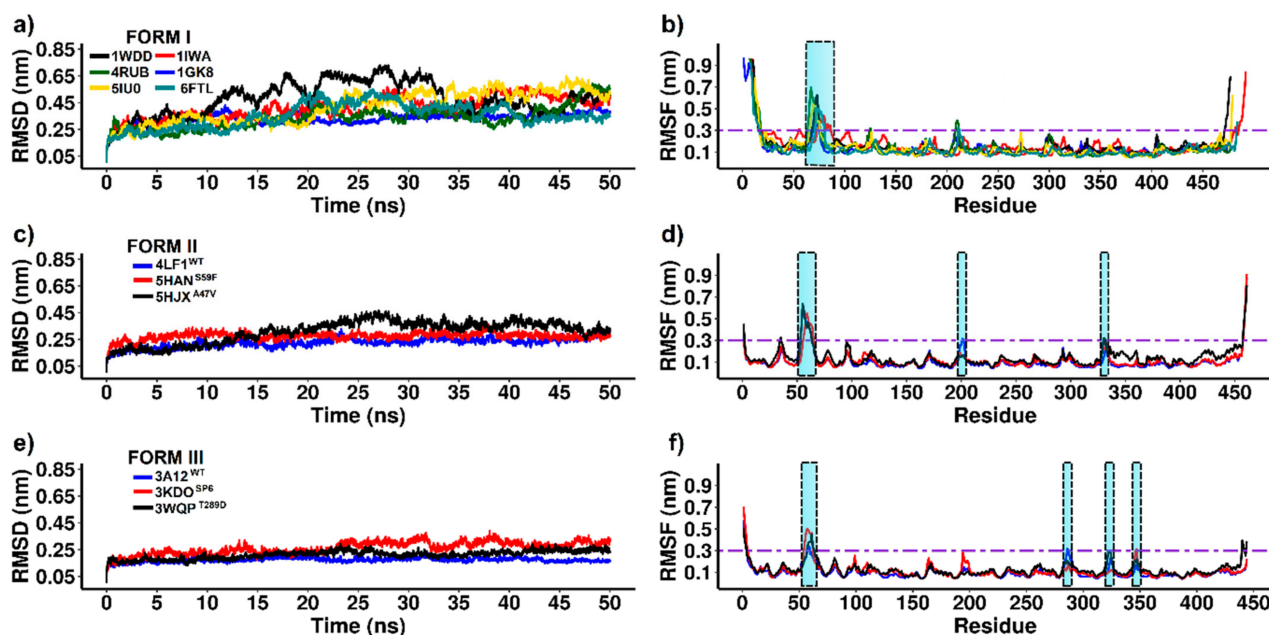


Figure 5. RMSD and RMSF profiles of RuBisCO forms I, II and III. (a) 50 ns RMSD of form I; (b) RuBisCO form I RMSF; (c) 50 ns RMSD of form II; (d) RuBisCO form II RMSF; (e) 50 ns RMSD of form III; (f) RuBisCO form III RMSF. RMSD was used to measure the deviations of the protein backbone from its original structural conformation to its final structural conformation. *C-alpha* atoms were used to calculate RMSF.

The RMSF of form I showed greater fluctuations in the N-terminal and C-terminal tails because the first and last amino acid residues form a loop-shaped structure. Likewise, the six systems showed greater flexibility in the loop located between α B and β C (≥ 0.3 nm) of the N-terminal domain spanning the following amino acid residues: 1WDD (Thr68–Ser76), 4RUB (Gly64–Thr75), 5IU0 (Trp66–Thr75), 1IWA (Ala73–Ala86), 1GK8 (Val69–Thr75), and 6FTL (Ser71–Thr80); see Table 1. On the other hand, the active site of RuBisCO (α/β barrel) was stable (RMSF ≤ 0.3 nm) because it is located in the TIM barrel domain that allows a protein to be slightly rigid (Figure 5b).

On the other hand, the analysis of form II included the 4LF1^{WT} system and two mutants, 5HAN^{S59F} and 5HJX^{A47V}. The results showed that 4LF1^{WT} (~0.24 nm ± 0.001) and the mutant 5HAN^{S59F} (~0.28 nm ± 0.002) had lower RMSD values and greater stability during the trajectories. Moreover, there was an increase in the RMSD value of the 5HJX^{A47V} mutant (~0.34 nm ± 0.002), and this can be attributed to the presence of valine 47 (A47V) in the α B region at the N-terminal end (Table 1). On the other hand, the 4LF1^{WT} structure and mutants (5HAN^{S59F} and 5HJX^{A47V}) had similar flexibility in most amino acid residues. Among the systems analyzed in form II, the greatest fluctuation was found in the loop between α B and β C (4LF1^{WT} (Gly53–Asp63 residues), 5HAN^{S59F} (Val56–Thr65 residues) and 5HJX^{A47V} (Thr54–Asp63 residues)), and other relevant fluctuations (≥ 0.3 nm) were: 4LF1^{WT} (Val201–Phe202, Pro458–Ala461 residues), 5HAN^{S59F} (Pro458–Ala461 residues) and 5HJX^{A47V} (Lys330–Met331, Pro458–Ala461 residues). In conclusion, the 5HAN^{S59F}

mutant analysis allowed key residues that reduced (Gly53–Glu57 residues, Val201–Phe202 residues) and increased (Phe59, Asp63–Phe64 residues) the flexibility by ≥ 0.1 nm to be identified (Table 1). Changes were found at loop 6 and in the loop between α B and β C, which is a critical region for gaseous substrate binding after RuBP enolization has been completed [18,63]. Moreover, the comparison between 4LF1^{WT} and 5HJX^{A47V} allowed the identification of key residues that increased (Gly35, Lys330–Met331) and reduced (Val201–Phe202) RuBisCO form II fluctuation by more than 0.1 nm (Table 1), which can be attributed to A47V mutation and which has an effect on RuBisCO catalytic activity [63].

Table 1. Average and standard error of RMSD in 12 RbcL structures. Regions were chosen according to their residues with the highest mean RMSF (≥ 0.3 nm).

Form	Protein	RMSD	Region ≥ 0.3 of RMSF	Sequence	Structures
I	1WDD ^{WT}	0.52 \pm 0.004	68–76	TVWTDGLTS	Loop connecting α B and β C
	4RUB ^{WT}	0.36 \pm 0.002	64–75; 125; 209–211	GTWTTVWTDGLT; F; QPF	Loop connecting α B and β C; α 0; Loop connecting β 2 and α 2
	5IU0 ^{WT}	0.46 \pm 0.003	22; 66–75	L; WTTVWTDGLT	N-terminal tail; Loop connecting α B and β C
	1IWA ^{WT}	0.44 \pm 0.003	55–56; 73–86; 482	PG; WTVVWTDLLTAA; T	β B; Loop connecting α B and β C; C-terminal tail
	1GK8 ^{WT}	0.34 \pm 0.002	69–75	VWTDGLT	Loop connecting α B and β C
	6FTL ^{WT}	0.37 \pm 0.002	71–80; 211–212	TVVWTDLLTA; NS	Loop connecting α B and β C; Loop connecting β 2 and α 2
II	4LF1 ^{WT}	0.24 \pm 0.001	53–63; 201–202	GTNVEVSTTDD; VF	Loop connecting α B and β C; Loop connecting β 2 and α 2
	5HAN ^{S59F}	0.28 \pm 0.002	56–65	VEVFTTDDFT	Loop connecting α B and β C
	5HJX ^{A47V}	0.34 \pm 0.002	54–63; 330–331	TNVEVSTTDD; KM	Loop connecting α B and β C; Loop 6
III	3A12 ^{WT}	0.18 \pm 0.001	58–59; 286	LY; A	Loop connecting α B and β C; α F
	3KDO ^{SP6}	0.28 \pm 0.002	55–62; 347	WTTLYPWY; N	Loop connecting α B and β C; Loop connecting α 6 and β 7
	3WQP ^{T289D}	0.22 \pm 0.001	57–63; 322	TLYPWYE; K	Loop connecting α B and β C; Loop 6

The RMSD analysis of form III showed that the 3A12^{WT} protein (~ 0.18 nm \pm 0.001) was more stable than 3KDO^{SP6} (~ 0.28 nm \pm 0.001) and 3WQP^{T289D} (~ 0.22 nm \pm 0.001); see Table 1. The 3KDO^{SP6} mutant showed regions where RMSF differed markedly from WT (Figure 5e). The residues Trp55–Tyr62 (loop that connects α B and β C) and Asn347 (loop that connects α 6 and β 7) showed, on average, a higher RMSF than WT (Table 1). However, the region that exhibited a lower RMSF in the 3KDO^{SP6} mutant with respect to WT was the α F region (Ala286); see Table 1. Moreover, the analysis of the 3WQP^{T289D} mutant showed a fluctuation greater than 0.1 nm in residues 60–61 (loop connecting α B and β C) and 322 (loop 6); see Table 1. Likewise, amino acid 322 is involved in direct interaction with the ligand CAP 2-carboxyarabinitol-1,5-diphosphate (C₆H₁₄O₁₃P₂). Thus, loop 6 is a region that plays a critical role in improving the enzymatic activity in the 3WQP^{T289D} mutant. Finally, our results of NMA and RMSF are in agreement because they were able to identify similar regions of greater flexibility in RuBisCO isoforms, where the loop between α B and β C presented greater flexibility.

3.3. Principal Component Analysis (PCA)

To obtain information on the conformational states of RbcL form I (5IU0, 1IWA, 1GK8, 1WDD, 4RUB and 6FTL), the PCA of the C α atoms was carried out. The first two PCs (PC 1/2) were taken in account. Figure 6 indicates the variance in the conformational distributions of proteins, where the display of continuous color points (from blue to white and to red) highlights periodic jumps between structural conformations. Moreover, the PC

1/2 of the MD trajectories were quite varied for the six systems, showing differences in the movement and stability of RuBisCO form I (Figure 6). In 5IU0, 1IWA, 1GK8 and 1WDD systems, there was greater correlated movement along the first two components, with a percentage of 85.5%, 80.4%, 76.7% and 75.6%, respectively, while in the 4RUB and 6FTL systems, the PC values were 70.6% and 70.5%, respectively (Figure 6). On the other hand, the PC 1/2 for 5IU0, 1GK8, 1IWA and 6FTL systems clearly shows the thermodynamically distinct periodic jumps (Figure 6), where most of the blue and red dots were assembled and distributed in opposite regions; therefore, proteins were in a relatively stable state in the system (Figure 6). Thus, the RbcL structures of *Oryza sativa* (1WDD) and *Nicotiana tabacum* (4RUB) showed a uniform distribution, overlapping PC subspace where there were not energy barriers, because most dots were in a scattered state. PCA analysis could suggest that the IB substructure of RbcL may undergo a periodic change in its conformation to reorient its domains (N-terminal and α/β barrel).

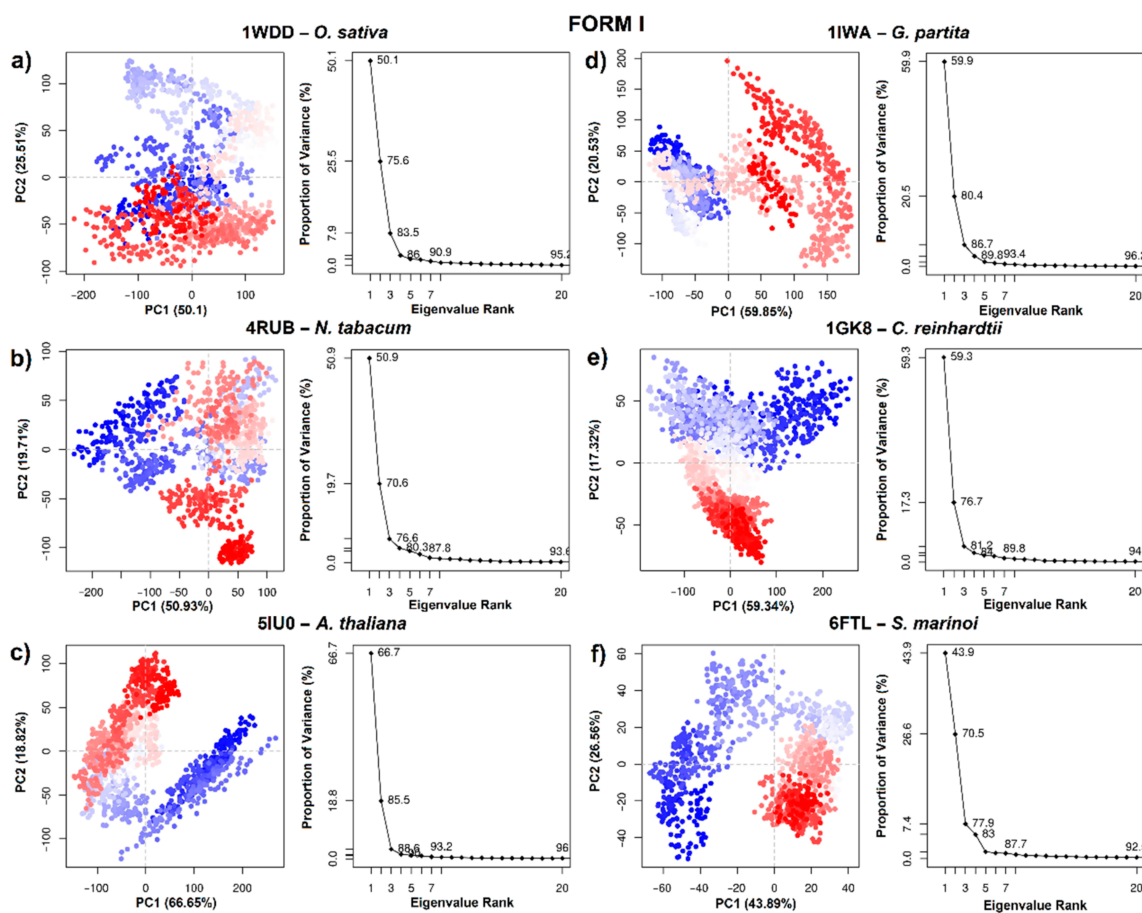


Figure 6. Principal component analysis of RuBisCO isoform I. (a) *O. sativa* (1WDD); (b) *N. tabacum* (4RUB); (c) *A. thaliana* (5IU0); (d) *G. partita* (1IWA); (e) *C. reinhardtii* (1GK8) and (f) *S. marinoi* (6FTL).

Regarding RuBisCO form II, PCA analysis allowed information on the conformational states of 4LF1^{WT} and two mutants (5HAN^{S59F} and 5HJX^{A47V}) to be obtained. The PC 1/2 of 4LF1^{WT}, 5HAN^{S59F} and 5HJX^{A47V} was 62.66%, 50.91% and 61.79%, respectively (Figure 7a–c). The scatter distribution of red and blue dots represents two different stable conformational states of the protein. The 4LF1^{WT} system was revealed to be more stable than the mutants (5HAN^{S59F} and 5HJX^{A47V}). Finally, the scatter plot of the 5HJX^{A47V} mutant (Figure 7c) showed the most unstable state of *R. palustris*. This is in agreement with RMSD results (Figure 5c), where 5HJX^{A47V} demonstrated more flexibility (~0.1 nm) than WT. Thus, the more dispersed conformational state was produced by repressor mutations (S59F and A47V).

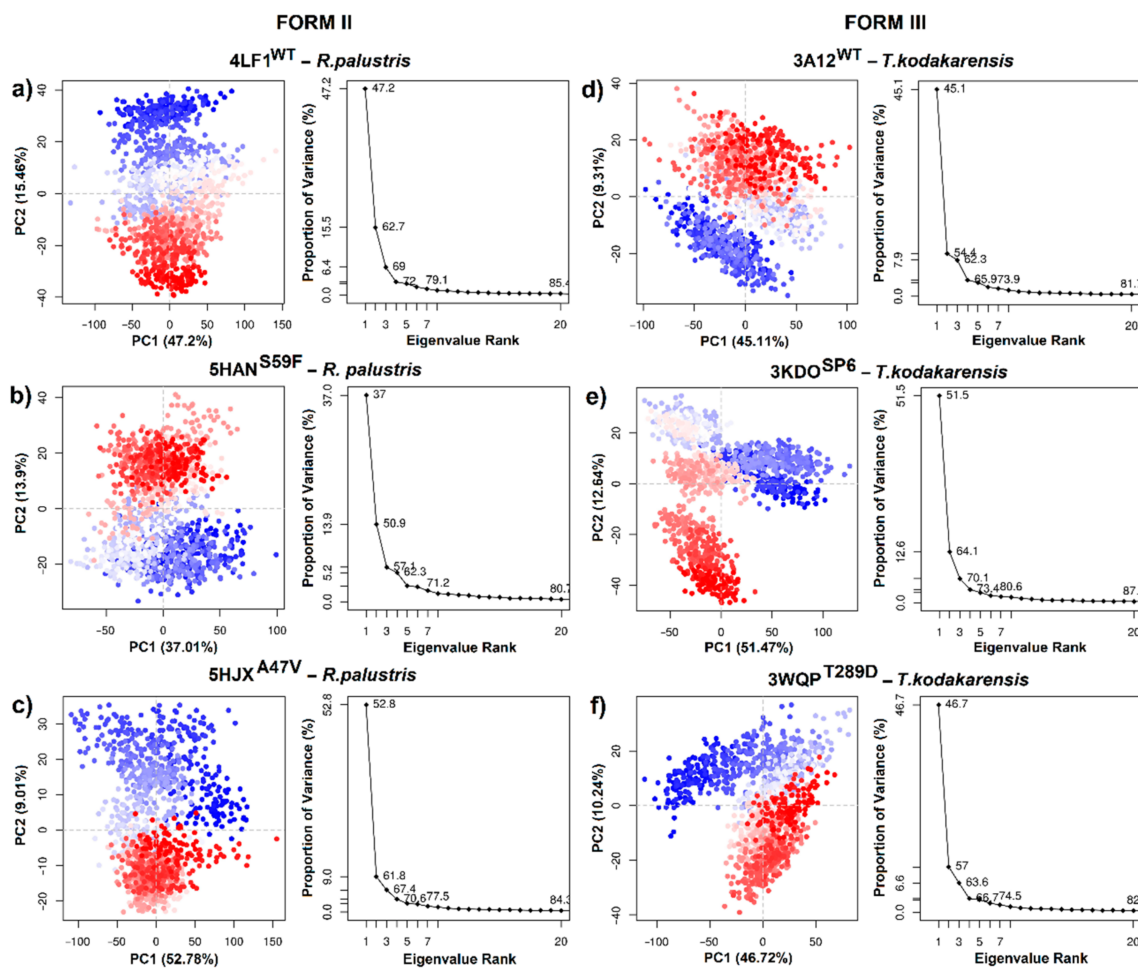


Figure 7. Principal component analysis of RuBisCO isoforms II (*R. palustris*) and III (*T. kodakarensis*). (a) 4LF1^{WT} (*R. palustris*); (b) 5HAN^{S59F} (*R. palustris* mutant); (c) 5HJX^{A47V} (*R. palustris* mutant); (d) 3A12^{WT} (*T. kodakarensis*); (e) 3KDO^{SP6} (*T. kodakarensis* mutant); and (f) 3WQP^{T289D} (*T. kodakarensis* mutant).

Figure 7 shows PCA analysis of *T. kodakarensis* with 1 wild-type (3A12^{WT}) and 2 mutants (3KDO^{SP6} and 3WQP^{T289D}). The first two eigenvectors captured most of the variance. The PCs (PC1/2) of the three systems contributed 54.42%, 64.11% and 56.96%, respectively (Figure 7d–f). Likewise, the analysis of the variance in the conformational distribution of 3KDO^{SP6} and 3WQP^{T289D} shows that mutants were energetically more stable than the WT system (Figure 7d–f). This analysis suggests that the WT may undergo a periodic change in its conformation to reorient its N-terminal or C-terminal domain. Differentiated grouping can be energy expensive; however, it can provide a control mechanism in the photosynthetic activity of RuBisCO.

3.4. Dynamic Cross-Correlation Matrix (DCCM)

One structural transition that is essential for the carboxylation of the 2,3-ene-diol(ate) intermediate is the closure of the active site of loop 6 in the large subunit and the concomitant movement of loop connecting α B and β C at the N-terminal end to stabilize the catalytic of loop 6 conformation [18,63,64]. Therefore, DCCM was performed to probe the conformational ensemble of these zones. Thus, similar correlation patterns were observed in 1WDD, 4RUB, 5IU0, 1GK8 and 6FTL (Figure 8a–d). This may provide insights into a conserved mechanism among *Chlamydomonas*, *Skeletonema*, *Arabidopsis*, tobacco and rice, since there was a correlated movement of residues 60–80 (part of α B region and the loop connecting α B and β C) at the N-terminal end with secondary structures as α 4, β 5, α F,

β F, α 5, β 6 and loop 6, which are located between residues 270–345 in the C-terminal end (Figure 8a–d). This correlation is important in RuBisCO since it connects the region with the greatest flexibility (loop connecting α B and β C) with the key substrate-binding residues H294, R295, H327 and K334. Moreover, the β C, loop CD, β D and α C structures had a strong negative correlation with α 4, β 5, α F, β F, α 5, β 6 and loop 6 structures. This suggests highly synchronized movements of the RuBisCO structure.

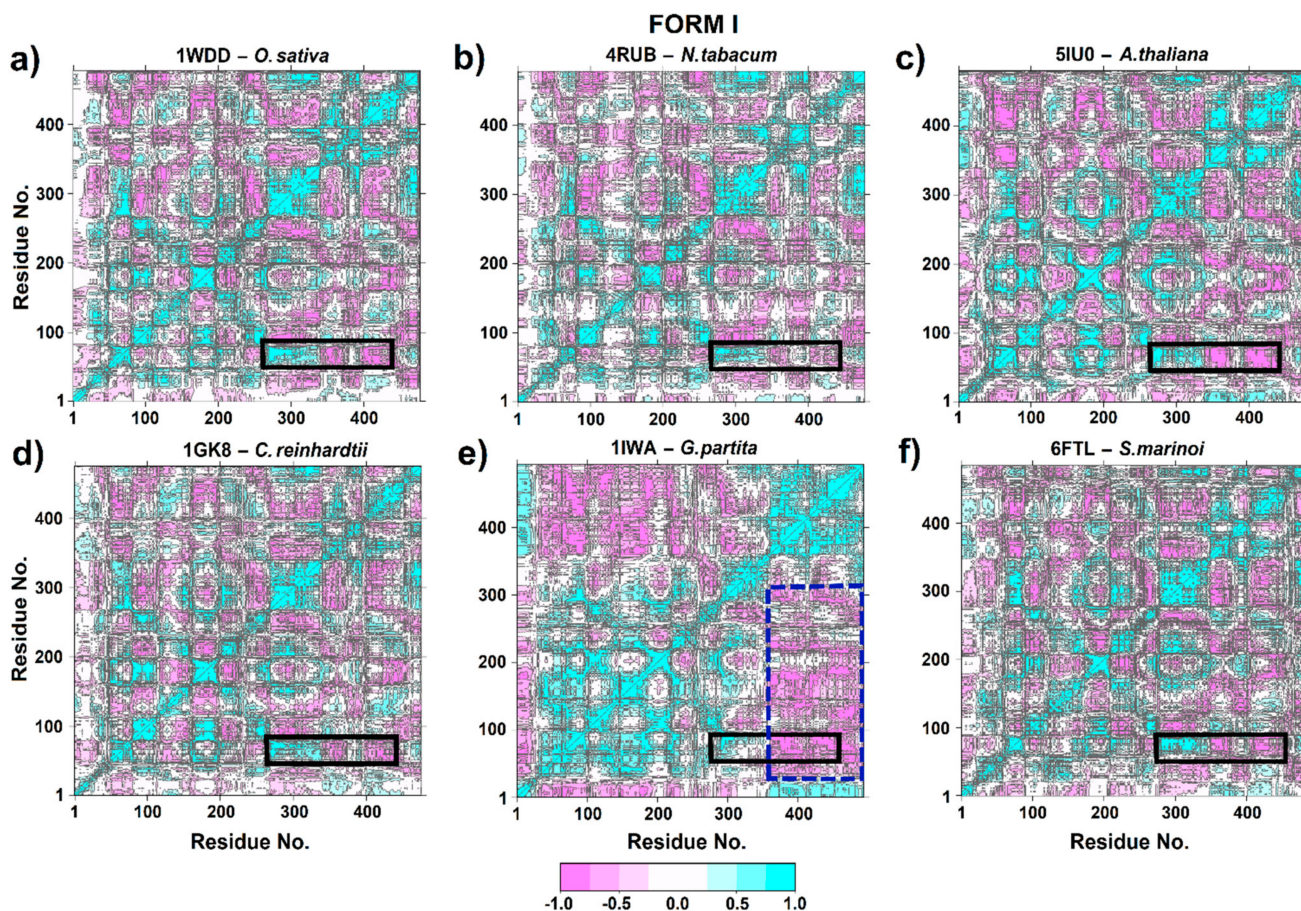


Figure 8. Cross-correlation analysis (DCCM) of RuBisCO form I. (a) *O. sativa* (1WDD); (b) *N. tabacum* (4RUB); (c) *A. thaliana* (5IU0); (d) *G. partita* (1IWA); (e) *C. reinhardtii* (1GK8) and (f) *S. marinoi* (6FTL). The color scale ranges from pink (for values ranging from -1 to -0.5) to white (-0.5 to 0.5) and to cyan (0.5 to 1).

On the other hand, *Galdieria partita* (1IWA) had a different correlation pattern than other isoform I proteins (Figure 8e). Thus, 1IWA presented a strong negative correlation (≥ -0.5) at residues 25–220 (secondary structures α -2, β B, α B, β C, loop CD, β D, α -1, α C, α 0, β E, α D and α E) with respect to amino acid residues located at position 350–493 (α 6, β 7, α 7, β 8, α 8, α G and α H) in the blue dotted line rectangle. However, a positive correlation was observed between the loop that connects α B and β C with the structures β 4, α 4, β 5, α F, β F, α 5, β 6 and loop 6 (black rectangle); see Figure 8e.

The DCCM plots for RuBisCO Form II showed that residues 53–63 exhibited anticorrelated movement with the structures β 4, α 4, β 5, α F, β F and α 5 (Ala255–Gly317). Moreover, the most flexible region (residues 53–63) located at the loop connecting α B and β C moves in the same direction (positive correlation > 0.5) as β 6, loop 6 and α 6 structures that are located between residues Gly326–Ala341 at the C-terminal domain (Figure 8a–c). This is because the movement of the N-terminal domain (the loop that connects α B and β C) towards the active site is important, since it is a key step in the catalytic mechanism of RuBisCO that involves CO_2 addition. Furthermore, the comparison between 5HAN^{S59F} (Figure 9b) and LF1^{WT} (Figure 9a) showed minor anticorrelations (blue dashed line box).

Therefore, S59F mutation is the most flexible region with a direct effect on loop 6 residues. Regarding the 5HJX^{A47V} mutant (Figure 9c), no significant changes were found in relation to 4LF1^{WT} (Figure 9a).

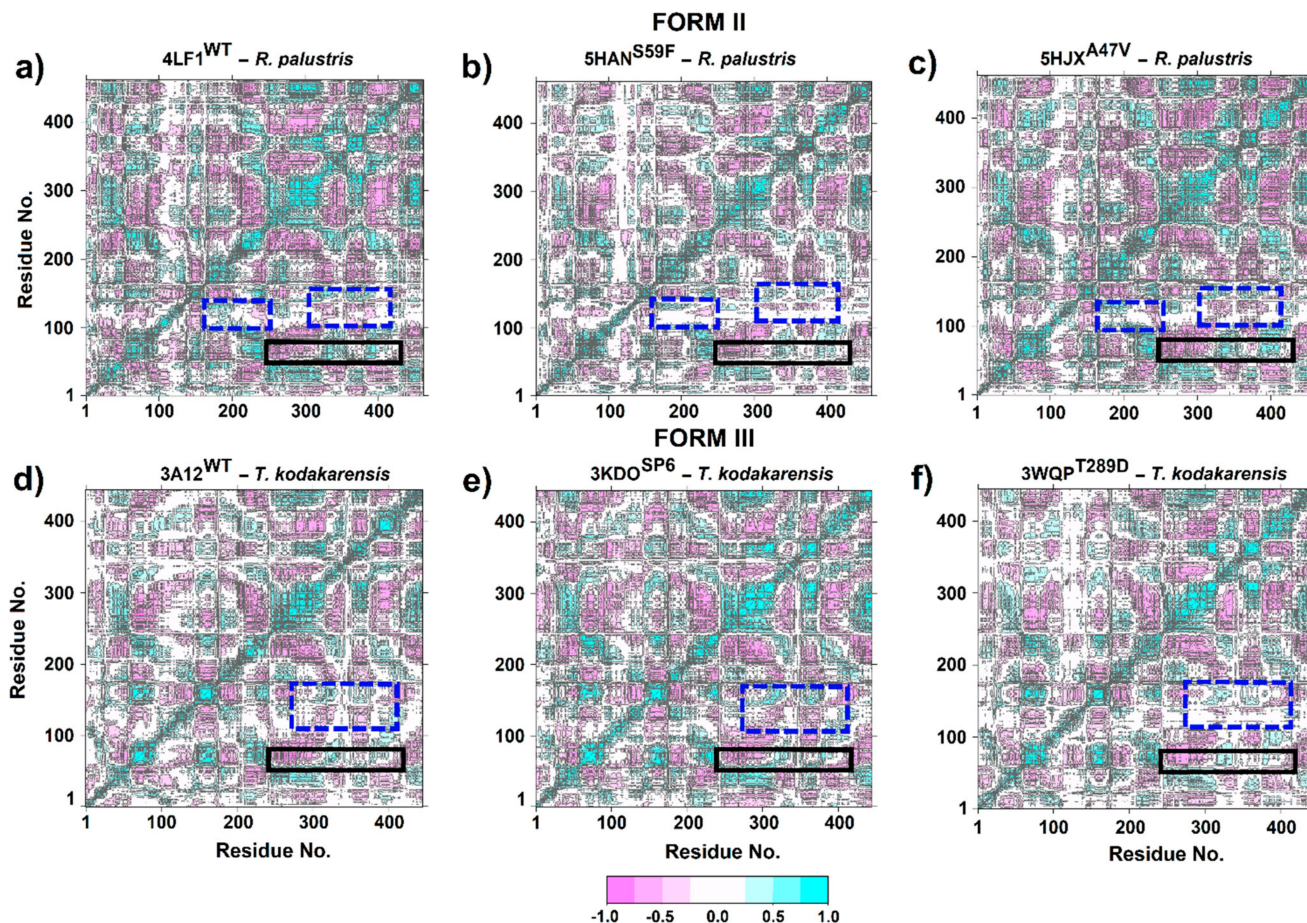


Figure 9. Cross-correlation analysis (DCCM) of RuBisCO form II (*R. palustris*) and III (*T. kodakarensis*). (a) 4LF1^{WT} (*R. palustris*); (b) 5HAN^{S59F} (*R. palustris* mutant); (c) 5HJX^{A47V} (*R. palustris* mutant); (d) 3A12^{WT} (*T. kodakarensis*); (e) 3KDO^{SP6} (*T. kodakarensis* mutant); and (f) 3WQP^{T289D} (*T. kodakarensis* mutant). The color scale ranges from pink (for values ranging from -1 to -0.5) to white (-0.5 to 0.5) and to cyan (0.5 to 1).

The DCCM analysis of the loop connecting α B and β C in RuBisCO form III demonstrated an anticorrelated movement (< -0.5) of Pro60-Ala72 residues with β 6 and loop 6 structures. However, Ser50-Tyr59 residues presented a positive correlation (> 0.5) with respect to β 6 and loop 6 structures (Figure 9). On the other hand, correlated movements were significantly reduced in 3WQP^{T289D} with respect to 3A12^{WT} (blue dotted line rectangle). This region comprises a positive correlation between α E, β 1 and loop structures that connect β 1 and α 1 with the β 6 and loop 6 structures. However, the positive correlation movements of these regions decreased in the 3WQP^{T289D} mutant (Figure 9f), since loop 6 (Lys322) mutation affects the correlation between specific residues, making them more flexible. Finally, the 3KDO^{SP6} mutant (Figure 9e) did not show visible changes in the residue-residue correlation patterns when compared to 3A12^{WT} (Figure 9d).

4. Discussion

RuBisCO plays a key role in carbon fixation on Earth. This enzyme possibly evolved during the Archean Eon [65] from an ancestral non-carbon fixing enzyme long before the appearance of the Calvin–Benson–Bassham cycle [3,66]. Thus, our phylogenetic analysis (Figure 2) allows the different RuBisCO isoforms to be identified, supporting the idea that photosynthetic RuBisCO (form I, II and III) and RubisCO-like protein (RLP, Form IV)

evolved from the same ancestral protein [67,68]. This would have allowed photosynthetic organisms to adopt different strategies to improve CO₂ specificity (as in the *Galdieria partita* case) [69], increase intracellular CO₂ concentration through a mechanism of carbon concentration [70], or inhabit ecological niches that have low levels of O₂/CO₂, such as in methanogenic organisms (as in the *Methanococcoides burtonii* case) [66].

The RbcL subunit is common among all isoforms (Figure 1). Currently, there are more than 64 RbcL structures in the PDB database (Figure 1, Table S1). Our PCA results using the Bio3D package [21] to identify three clusters with different structural flexibilities (Figure 1). This also allowed the identification of the key amino acid residues, several structural characteristics, and conformational changes that are critical for folding and catalytic activity (Figure 3). The largest cluster corresponded to Isoform I, which included higher plants, green algae, blue-green algae, cyanobacteria, diatoms, and proteobacteria (Figure 1). The emergence of form I complexes through the incorporation of small subunits represents a transitional key that is little understood in RuBisCO evolution [19]. However, the 6URA structure of *Promineofilum breve* (a bacteria) has a structural flexibility that allowed grouping of isoform I (Figure 1), taking into account that 6URA does not present small RbcS subunits [71] and also presents deletions in secondary structural elements such as loop-CD and the loop that connects $\alpha 8$ - αG , making it a reference point to advance our understanding of isoform I evolution. In the second cluster (Figure 1), proteobacteria of RuBisCO Form II were mostly reported, except for the methanogenic archaea *Methanococcoides burtonii* (5MAC), which presents an isoform II/III [5,65]. It also shows a unique insert of 26–30 amino acids between the $\alpha 6$ and $\beta 7$ secondary structures at the bottom of the $\beta \alpha$ -barrel [72]. Methanogens like *Methanococcoides burtonii* (5MAC) are strictly anaerobic and cannot survive in the presence of oxygen [73]. Thus, their RuBisCO are not under selection pressure to mitigate the competitive binding of O₂ over CO₂.

Despite having only 30% amino acid identity, the multiple sequence alignment analysis among RuBisCO isoforms showed large changes in the N-terminal and C-terminal regions (Figure 3). Moreover, the structures retain the residues involved in substrate binding and catalytic activity (Figure 3), supporting the idea that they are critical to folding and maintaining the overall structure and function of the photosynthetic RuBisCO [1,66]. Moreover, with the molecular dynamics analysis, it was possible to sample the transition of the RuBisCO conformation in 50 ns, and the RMSD results were consistent with previous studies [14,45,46]. Our results added more theoretical evidence on the structural movements of RuBisCO, helping to understand the structure flexibility and how this could affect the synchronization of the residues and the closing mechanism, which are still unknown [19]. Thus, the NMA and RMSF results revealed similar flexibility patterns in the RbcL structures (Figures 4 and 5), showing the distribution of temperature factors (B-factors) and the fact that loops which are more flexible during catalysis are also more flexible in crystallized structures [74–76]. Consequently, our results indicate that secondary structures such as the loop connecting αB and βC (~64–85) and the tails in the N-terminal and C-terminal region show greater fluctuations in the three isoforms (Figure 5), and the structural movements are related to the structural changes and activities of RuBisCO during the transition from its open to closed state [74,77]. Thus, closed-state carboxylation of RuBisCO is more likely when the substrate is attached, whereas fluctuations in larger tails can also cause structural changes and deactivations of RuBisCO.

According to DCCM analysis of isoform I, the structure of 1WDD, 4RUB, 5IU0, 1GK8 and 6FTL preserved the direction of the movements in the time of residues ~64–85 (part of the αB region and the connecting loop between αB and βC) and ~270–345 residues (related to the secondary structures $\alpha 4$, $\beta 5$, αF , βF , $\alpha 5$, $\beta 6$ and loop 6); see Figure 8. This correlation is important for RuBisCO activity, since it connects the region of greatest flexibility (connecting loop between αB and βC) with the key residues (H294, R295, H327 and K334) that bind the substrate (Figure 4), where the two states (open and closed) of RuBisCO are distinguished by the degree of accessibility of the solvent to the active site [74]. The closed state is associated with substrates and inhibitors (CABP 2-carboxyarabinitol

1,5-bisphosphate) that are attached to the active site. This would be achieved through a movement of loop 6 (residues 331–338) and the connecting loop between α B and β C (residues ~64–85); see Table 1. In addition, the N-terminal and C-terminal loops function like latches that hold loop ~64–85 and loop 6 in their closed positions with an extremely slow release of CABP [74]. Moreover, RuBisCO studies carried out on spinach and wheat were able to show that residues ~8 to 20 at the N-terminal end are only ordered when the active site is closed [78]. In the closed conformation, the N-terminal end (Phe13, Lys14, Gly16 and Lys18) is placed directly on the connecting loop between α B and β C (~64–85 residues), which coordinates the P1 site of the substrate [78]. In contrast, the open state is associated with weak products being attached (metal ions that are catalytically inert). In the open state, loop 6 of the α/β barrel is away from the active site, and the C-terminal end of the large subunit is disordered, so the active site is open.

On the other hand, *Galdieria partita* is a thermophilic red alga with a high specificity, and this alga showed a marked difference with respect to the RuBisCO of higher plants [79]. Our results are in agreement with Watanabe et al. [79] because the movements of the RbcL structure of *G. partita* presented a different residue-residue correlation pattern than IB and IC/D structures (Figure 8e). Although *G. partita* has more residues in the N-terminal and C-terminal regions (Figure 3), this phenomenon could play an important role in the structural movement of the RuBisCO enzyme, as indicated by some studies [78]. Likewise, the RMSF analysis of 1IWA (*G. partita*) showed the greatest flexibility between residues ~Trp73–Ala86 (the connecting loop between α B and β C). Moreover, it is necessary to point that Thr74 in the RbcL sequence of *G. partita* is an evolutionarily conserved amino acid that binds the phosphate groups of RuBP. Thus, the Thr74 residue can alter the packing in the C-terminal end [79]. Likewise, when RuBisCO is in the open state, the hydrogen bond breaks between P1 and Thr74, as well as between Thr76 and Trp470 [79]. Thr74 is stabilized by the carbonyl of the Thr76 backbone, while the stabilization of the Thr76 and Trp470 side chains occurs in the closed state of loop 6 [79]. Moreover, the work of Satagopan et al. [63] on *R. rubrum* (RuBisCO isoform II) is in agreement with Watanabe et al. [79], because the link between the α B and β C loop with loop 6 was also identified; this link allows the conformation of the catalytic loop to be stabilized. Satagopan et al. [63] used mutants (5HAN^{S59F} and 5HJX^{A47V}) in *R. rubrum* where CO₂ was the only carbon source; their results showed a similar biological growth between mutants and a decrease with respect to WT [63,80]. Thus, our PCA results of DM indicated that mutants (5HAN^{S59F} and 5HJX^{A47V}) would have undergone a change from an unstable intermediate conformation (4LF1^{WT}) to an unstable one, showing a greater variation in the conformational distribution of the RuBisCO (Figure 7b,c), being consistent with the RMSD results where 5HJX^{A47V} was more flexible (~0.1 nm) than WT (Figure 5). Consequently, the fluctuation in the 5HJX^{A47V} mutant was increased by more than 0.1 nm in the key residues (Lys330–Met331) that are located in loop 6 (Table 1). Some experimental studies showed that the mutations (M331L and M331A) affected the growth of the purple photosynthetic bacterium *R. palustris*.

Thus, the residue Met331 and its interactions seem to be specific and critical for the addition of CO₂ to the intermediate 2,3-enediol(ate) derived from RuBP [63,80]. On the contrary, strains with double mutations (A47V/M331A and S59F/M331A) in *R. palustris* did show growth. Thus, growth inhibition induced by M331A was suppressed by the substitution of A47V and S59F [63]. To explain this phenomenon, Satagopan et al. [80] evaluated the movements and crystallized structures of WT and mutants, where the α carbon of Ala47 (structure α B) was ~15 Å away from Met331 (loop 6) [80]. Furthermore, our results indicate that Ala47 and Met331 move in the same direction (positive correlation > 0.5; Figure 9a–c). Likewise, the side chain is found in a hydrophobic environment within 4 Å of Ala70 and Val72 [80]. Substitution with a Val (A47V) would push the α B structure towards the active site, and thus Lys330 and Met331 showed greater flexibility (Table 1; Figure 5). Glu49 located in the α B structure works to stabilize Lys330 and helps to close loop 6 during catalysis, and Thr54 binds phosphate P1 of the substrate [78,81,82].

Thermococcus kodakarensis is a hyperthermophilic archaea. Its optimal growth is at 85 °C, and it can exceed the activity of RuBisCO spinach by 20 times [12,83], but at room temperature its activity is only one-eighth [12]. For this reason, it is necessary to develop new *T. kodakarensis* strains with good photosynthetic performance at room temperatures. In this sense, 3WQP^{T289D} [12] and 3KDO^{SP6} [10] mutants were developed. Their results show an increase in the carboxylase activity of 24% and 31%, respectively. In the 3KDO^{SP6} mutant, residues of the $\alpha 6$ region were replaced by 11 amino acid residues from spinach (E326-L336 “ERDITLGFVDL”). Consequently, Nishitani et al. [10] evaluated the flexibility between 3A12^{WT} and 3KDO^{SP6}, showing that there is an increase in temperature factors (B-Factors Å²) in the secondary structure $\alpha 6$. Moreover, our ANM results showed a greater fluctuation in the loop between αB and βC and in the loop that connects $\alpha 6$ and $\beta 7$ (His341-Ala361) with respect to isoforms I and II. Likewise, the mutants were more energetically stable than WT (Figure 7e,f), suggesting that WT may undergo a periodic change in its conformation to reorient its N-terminal or C-terminal domain, since the differentiated clustering of RuBisCO conformational distributions can be energetically expensive [84]. However, it can also provide a control mechanism in the photosynthetic activity of RuBisCO.

In addition, 3KDO^{SP6} showed greater flexibility of ~0.15 nm in the loop that connects αB and βC (Figure 5f) and the 3WQP^{T289D} mutant in the residue Lys322 (a residue of loop 6 which is catalytically critical) (Figure 5f, Table 1). From this analysis, the changes between αF and βF in 3WQP^{T289D} can influence loop 6's movement (Table 1). Therefore, an increase in the loop flexibility between αB and βC , Lys322 or residues in the vicinity of the catalytic center is important to increase the catalytic activity of RuBisCO from *T. kodakarensis* at room temperature. Finally, our MD simulation results corroborate very well the work of Satagopan et al. [63] and Fujihashi et al. [12]. Likewise, the regions of greater flexibility and active sites exhibit highly correlated or anticorrelated movements between the different isoforms (Figures 8 and 9), building a dynamic correlation network where information signal are transmitted [48]. Therefore, it is necessary to build next-generation computational tools, where a PAN-DM approach will allow us to integrate very complex information on molecular structure, dynamics and evolution.

5. Conclusions

Our PCA showed a wide range of the conformational space of the RuBisCO crystal structures, allowing the identification of different isoforms. Likewise, phylogenetic analysis supports the idea that RuBisCO evolved from the same ancestral enzyme, conserving the residues involved in substrate binding and catalytic activity. On the other hand, molecular dynamics analyses were able to sample the transition of RuBisCO conformation in 50 ns. Thus, the NMA and RMSF results revealed similar flexibility patterns. The places where the secondary structures loop between αB and βC as well as the tails in the N-terminal and C-terminal regions show greater fluctuations among the three isoforms, and their movements are possibly related to the structural changes and functional activities of the RuBisCO enzyme during the transition from its open to closed state. On the other hand, the DCCM results indicate that there are changes in the movement direction of the secondary structures of the three isoforms. However, movements in the same direction are preserved in loop 6 and the connecting loop between αB and βC . This correlation is important for enzymatic activity and to stabilize the conformation of the catalytic loop. In isoform I, the 1WDD, 4RUB, 5IU0, 1GK8 and 6FTL structures showed a positive correlation between the residue movement of ~64–85 (part of the αB region and the connecting loop between αB and βC) and residues ~270–345 (secondary structures $\alpha 4$, $\beta 5$, αF , βF , $\alpha 5$, $\beta 6$ and loop 6). On the other hand, 1IWA (*Galdieria partita*) had a marked difference in the direction of structural movements with respect to the others' isoform I. This could have a key role, and it is required to deepen its study with mutants. On the other hand, the PCA results of *R. rubrum* (RuBisCO form II) indicated that mutants (5HAN^{S59F} and 5HJX^{A47V}) would have undergone a change from an unstable intermediate conformation (4LF1^{WT}) to an unstable one, showing a greater variation in the conformational distribution of RuBisCO.

Consequently, the 5HJX^{A47V} mutant allowed the fluctuation of the key residues Lys330–Met331 in loop 6 to be increased by more than 0.1 nm. Regarding isoform III, the mutants (3KDO^{SP6}, 3WQP^{T289D}) of the hyperthermophilic archaea *Thermococcus kodakarensis* were more energetically stable than WT, suggesting that the WT may undergo a periodic change in its conformation to reorient its N-terminal or C-terminal domain, a control mechanism in the photosynthetic activity of RuBisCO. 3KDO^{SP6} showed a greater flexibility of ~0.15 nm in the loop between α B and β C and Lys322 in the 3WQP^{T289D} mutant (a catalytic residue critical in loop 6) with respect to 3A12^{WT}. Thus, an increase in the loop flexibility between α B and β C, Lys322 or residues in the vicinity of the catalytic center is important to increase the photosynthetic efficiency of RuBisCO from *T. kodakarensis* at room temperature. Finally, our results added more evidence regarding the structural movements of RuBisCO, helping to understand the details of the synchronization and the closing mechanism that are still unknown.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/biom11121761/s1>, Table S1: 46 structures of isoforms I, II and III from wild-type RbcL and mutants in proteobacteria and archaea by X-ray crystallography, Figure S1: Analysis Pipeline of RuBisCO using Bio3d package.

Author Contributions: Conceptualization, V.C.; data curation, V.C.; formal analysis, V.C.; funding acquisition, G.Z.; investigation, V.C. and G.Z.; methodology, V.C.; project administration, V.C. and G.Z.; software, V.C.; supervision, G.Z.; validation, V.C.; visualization, V.C.; writing—original draft, V.C. and G.Z.; writing—review & editing, V.C. and G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by PROCENCIA grant numbers 177-2015-FONDECYT and 159-2018-FONDECYT-BM-IADT-AV.

Institutional Review Board Statement: This study did not involve humans or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and/or analyzed during the current study, are available on request from the corresponding author.

Acknowledgments: The authors want to thank: Emilio Paredes Solis for helping to install the Gromacs software, Georcki Rojon-Palacios for sharing his knowledge in computational biophysics and Haoran Yu for sharing codes for molecular dynamics. We also thank two anonymous reviewers for providing constructive and helpful comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Andersson, I.; Backlund, A. Structure and Function of Rubisco. *Plant Phys. Biochem.* **2008**, *46*, 275–291. [[CrossRef](#)]
2. Tabita, F.R.; Hanson, T.E.; Li, H.; Satagopan, S.; Singh, J.; Chan, S. Function, Structure, and Evolution of the RubisCO-Like Proteins and Their RubisCO Homologs. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 576–599. [[CrossRef](#)] [[PubMed](#)]
3. Erb, T.; Zarzycki, J. A Short History of RubisCO: The Rise and Fall (?) of Nature's Predominant CO₂ Fixing Enzyme. *Plant Biotechnol.* **2018**, *49*, 100–107. [[CrossRef](#)]
4. Stec, B. Structural Mechanism of RuBisCO Activation by Carbamylation of the Active Site Lysine. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18785–18790. [[CrossRef](#)] [[PubMed](#)]
5. Liu, D.; Chettiyan, R.; Ramya, S.; Mueller-cajar, O. Surveying the Expanding Prokaryotic Rubisco Multiverse. *FEMS Microbiol. Lett.* **2017**, *1–10*. [[CrossRef](#)] [[PubMed](#)]
6. Kitano, K.; Maeda, N.; Fukui, T.; Atomi, H.; Imanaka, T.; Miki, K. Crystal Structure of a Novel-Type Archaeal Rubisco with Pentagonal Symmetry. *Structure* **2001**, *9*, 473–481. [[CrossRef](#)]
7. Kacar, B.; Hanson-smith, V.; Adam, Z.; Boekelheide, N. Constraining the Timing of the Great Oxidation Event within the Rubisco Phylogenetic Tree. *Geobiology* **2017**, *No. May*, 628–640. [[CrossRef](#)]
8. Mueller-cajar, O.; Morell, M.; Whitney, S.M. Directed Evolution of Rubisco in *Escherichia Coli* Reveals a Specificity-Determining Hydrogen Bond in the Form II Enzyme. *Biochemistry* **2007**, *46*, 14067–14074. [[CrossRef](#)]

9. Maeda, N.; Kitano, K.; Fukui, T.; Ezaki, S.; Atomi, H.; Miki, K.; Imanaka, T. Ribulose Bisphosphate Carboxylase/Oxygenase from the Hyperthermophilic Archaeon *Pyrococcus Kodakaraensis* KOD1 Is Composed Solely of Large Subunits and Forms a Pentagonal Structure. *J. Mol. Biol.* **1999**, *293*, 57–66. [[CrossRef](#)]
10. Nishitani, Y.; Yoshida, S.; Fujihashi, M.; Kitagawa, K.; Doi, T.; Atomi, H.; Imanaka, T.; Miki, K. Structure-Based Catalytic Optimization of a Type III Rubisco from a Hyperthermophile. *J. Biol. Chem.* **2010**, *285*, 39339–39347. [[CrossRef](#)]
11. Yoshida, S.; Atomi, H.; Imanaka, T. Engineering of a Type III Rubisco from a Hyperthermophilic Archaeon in Order to Enhance Catalytic Performance in Mesophilic Host Cells. *Appl. Environ. Microbiol.* **2007**, *73*, 6254–6261. [[CrossRef](#)]
12. Fujihashi, M.; Nishitani, Y.; Kiriya, T.; Aono, R.; Sato, T.; Takai, T.; Tagashira, K.; Fukuda, W.; Atomi, H.; Imanaka, T.; et al. Mutation Design of a Thermophilic Rubisco Based on Three-Dimensional Structure Enhances Its Activity at Ambient Temperature. *Prot. Struct. Funct. Bioinform.* **2016**, *84*, 1339–1346. [[CrossRef](#)] [[PubMed](#)]
13. Yu, H.; Dalby, P.A. Coupled Molecular Dynamics Mediate Long- and Short-Range Epistasis between Mutations That Affect Stability and Aggregation Kinetics. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11043–E11052. [[CrossRef](#)]
14. Faulkner, M.; Szabó, I.; Weetman, S.L.; Sicard, F.; Huber, R.G.; Bond, P.J.; Rosta, E.; Liu, L.N. Molecular Simulations Unravel the Molecular Principles That Mediate Selective Permeability of Carboxysome Shell Protein. *Sci. Rep.* **2020**, *10*, 1. [[CrossRef](#)]
15. Tabita, F.R.; Hanson, T.E.; Satagopan, S.; Witte, B.H.; Kreel, N.E. Phylogenetic and Evolutionary Relationships of RubisCO and the RubisCO-like Proteins and the Functional Lessons Provided by Diverse Molecular Forms. *R. Soc.* **2008**, 2629–2640. [[CrossRef](#)] [[PubMed](#)]
16. Tabita, F.R.; Satagopan, S.; Hanson, T.E.; Kreel, N.E.; Scott, S.S. Distinct Form I, II, III, and IV Rubisco Proteins from the Three Kingdoms of Life Provide Clues about Rubisco Evolution and Structure/Function Relationships. *J. Exp. Botany* **2008**, *59*, 1515–1524. [[CrossRef](#)]
17. Li, H.; Sawaya, M.R.; Tabita, F.R.; Eisenberg, D. Crystal Structure of a RuBisCO-like Protein from the Green Sulfur Bacterium *Chlorobium Tepidum*. *Structure* **2005**, *13*, 779–789. [[CrossRef](#)] [[PubMed](#)]
18. Duff, A.P.; Andrews, T.J.; Curmi, P.M.G. The Transition between the Open and Closed States of Rubisco Is Triggered by the Inter-Phosphate Distance of the Bound Bisphosphate. *J. Mol. Biol.* **2000**, *298*, 903–916. [[CrossRef](#)]
19. Genkov, T.; Spreitzer, R.J. Highly Conserved Small Subunit Residues Influence Rubisco Large Subunit Catalysis. *J. Biol. Chem.* **2009**, *284*, 30105–30112. [[CrossRef](#)]
20. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissing, H.; Shindyalov, I.; Bourne, P. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
21. Skjaerven, L.; Yao, X.Q.; Scarabelli, G.; Grant, B.J. Integrating Protein Structural Dynamics and Evolutionary Analysis with Bio3D. *BMC Bioinform.* **2014**, *15*, 1–11. [[CrossRef](#)] [[PubMed](#)]
22. Grant, B.J.; Skjaerven, L.; Yao, X.Q. The Bio3D Packages for Structural Bioinformatics. *Prot. Sci.* **2021**, *30*, 20–30. [[CrossRef](#)] [[PubMed](#)]
23. Bakan, A.; Meireles, L.M.; Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011**, *27*, 1575–1577. [[CrossRef](#)] [[PubMed](#)]
24. Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T.R.; Bahar, I. Structural Bioinformatics Evol and ProDy for Bridging Protein Sequence Evolution and Structural Dynamics. *Bioinformatics* **2014**, *30*, 2681–2683. [[CrossRef](#)] [[PubMed](#)]
25. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
26. Kalenkiewicz, A.; Grant, B.J.; Yang, C.Y. Enrichment of Druggable Conformations from Apo Protein Structures Using Cosolvent-Accelerated Molecular Dynamics. *Biology* **2015**, *4*, 344–366. [[CrossRef](#)]
27. Edgar, R.C. MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity. *BMC Bioinform.* **2004**, *5*, 1–19. [[CrossRef](#)] [[PubMed](#)]
28. Blow, D. *Outline of Crystallography for Biologists*, 1st ed.; Oxford University Press: New York, NY, USA, 2002.
29. Yu-Feng, H. *Study of Mining Protein Structural Properties and Its Application*; National Taiwan University: Taipei, Taiwan, 2007.
30. Hanson-Smith, V.; Johnson, A. PhyloBot: A Web Portal for Automated Phylogenetics, Ancestral Sequence Reconstruction, and Exploration of Mutational Trajectories. *PLoS Comput. Biol.* **2016**, *12*, 1–10. [[CrossRef](#)] [[PubMed](#)]
31. Liu, Y.; Schmidt, B.; Maskell, D.L. MSAProbs: Multiple Sequence Alignment Based on Pair Hidden Markov Models and Partition Function Posterior Probabilities. *Bioinformatics* **2010**, *26*, 1958–1964. [[CrossRef](#)]
32. Le, S.Q.; Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Mol. Biol. Evolution.* **2008**, *25*, 1307–1320. [[CrossRef](#)]
33. Lartillot, N.; Philippe, H. A Bayesian Mixture Model for across-Site Heterogeneities in the Amino-Acid Replacement Process. *Mol. Biol. Evol.* **2004**, *21*, 1095–1109. [[CrossRef](#)]
34. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* **2019**, *35*, 4453–4455. [[CrossRef](#)] [[PubMed](#)]
35. Guindon, S.; Dufayard, J.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
36. Tamura, K.; Stecher, G.; Peterson, D.; Filipowski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)]
37. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinform.* **2016**, *54*, 1–37. [[CrossRef](#)]

38. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Prot. Sci.* **2018**, *27*, 293–315. [[CrossRef](#)] [[PubMed](#)]
39. Luthy, R.; Bowie, J.; Eisenberg, D. Assessment of Protein Models with Three-Dimensional Profiles. *Nature.* **1992**, *359*, 83–85. [[CrossRef](#)] [[PubMed](#)]
40. Bowie, J.; Luthy, R.; Eisenberg, D. A Method to Identify Protein Sequences That Fold into a Known Three-Dimensional Structure. *Science* **1991**, *253*, 164–169. [[CrossRef](#)]
41. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [[CrossRef](#)]
42. Huang, J.; MacKerell, A. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *30*, 2135–2145. [[CrossRef](#)]
43. Ahrari, S.; Khosravi, F.; Osouli, A.; Sakhteman, A.; Nematollahi, A.; Ghasemi, Y.; Savardashtaki, A. MARK4 Protein Can Explore the Active-like Conformations in Its Non-Phosphorylated State. *Sci. Rep.* **2019**, *9*, 1–14. [[CrossRef](#)] [[PubMed](#)]
44. Salmas, R.E.; Unlu, A.; Yurtsever, M.; Noskov, S.Y.; Durdagi, S. In Silico Investigation of PARP-1 Catalytic Domains in Holo and Apo States for the Design of High-Affinity PARP-1 Inhibitors. *J. Enzym. Inhib. Med. Chem.* **2016**, *31*, 112–120. [[CrossRef](#)]
45. Guinot, A.D.M. Structural Studies of Different Form I Rubiscos Using Molecular Dynamics Simulations. Doctoral Dissertation, Imperial College London, London, UK, 2016. [[CrossRef](#)]
46. Siqueira, A.S.; Lima, A.R.J.; Dall’Agnol, L.T.; de Azevedo, J.S.N.; da Silva Gonçalves Vianez, J.L.; Gonçalves, E.C. Comparative Modeling and Molecular Dynamics Suggest High Carboxylase Activity of the Cyanobium Sp. CACIAM14 RbcL Protein. *J. Mol. Model.* **2016**, *22*, 3. [[CrossRef](#)]
47. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472. [[CrossRef](#)]
48. Yu, H.; Dalby, P.A. A Beginner’s Guide to Molecular Dynamics Simulations and the Identification of Cross-Correlation Networks for Enzyme Engineering. *Methods Enzymol.* **2020**, *643*, 15–49. [[CrossRef](#)] [[PubMed](#)]
49. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Grap.* **1996**, *14*, 33–38. [[CrossRef](#)]
50. Schrodinger, L. The PyMOL Molecular Graphics System, Version 1.3r1. 2010. Available online: <https://www.mdpi.com/1422-0067/21/19/7166/htm> (accessed on 23 September 2021).
51. Martínez, L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. *PLoS ONE* **2015**, *10*, 1–10. [[CrossRef](#)]
52. Lesgidou, N.; Eliopoulos, E.; Goulielmos, G.; Vlassi, M. Insights on the Alteration of Functionality of a Tyrosine Kinase 2 Variant: A Molecular Dynamics Study. *Bioinformatics* **2018**, *34*, i781–i786. [[CrossRef](#)]
53. Hong, L.; Ying, M.; Zheng, C.-J.; Jin, W.-Y.; Liu, W.-S.; Wang, R.-L. Exploring the Effect of D61G Mutation on SHP2 Cause Gain of Function Activity by a Molecular Dynamics Study. *J. Biomol. Struct. Dyn.* **2018**, *36*, 3856–3868. [[CrossRef](#)]
54. Rajapaksha, H.; Pandithavidana, D.; Dahanayake, J. Demystifying Chronic Kidney Disease of Unknown Etiology (CKDu): Computational Interaction Analysis of Pesticides and Metabolites with Vital Renal Enzymes. *Biomolecules* **2021**, *11*, 261. [[CrossRef](#)]
55. Zalewski, M.; Kmiecik, S.; Kolinski, M. Molecular Dynamics Scoring of Protein–Peptide Models Derived from Coarse-Grained Docking. *Molecules* **2021**, *26*, 3293. [[CrossRef](#)]
56. Bahar, I.; Atilgan, A.R.; Erman, B. Direct Evaluation of Thermal Fluctuations in Proteins Using a Single-Parameter Harmonic Potential. *Fold. Design* **1997**, *2*, 173–181. [[CrossRef](#)]
57. Wang, R.R.; Ma, Y.; Du, S.; Li, W.Y.; Sun, Y.Z.; Zhou, H.; Wang, R.L. Exploring the Reason for Increased Activity of SHP2 Caused by D61Y Mutation through Molecular Dynamics. *Comput. Biol. Chem.* **2019**, *78*, 133–143. [[CrossRef](#)]
58. Li, W.Y.; Wei, H.Y.; Sun, Y.Z.; Zhou, H.; Ma, Y.; Wang, R.L. Exploring the Effect of E76K Mutation on SHP2 Cause Gain-of-Function Activity by a Molecular Dynamics Study. *J. Cell. Biochem.* **2018**, *119*, 9941–9956. [[CrossRef](#)]
59. Ichiye, T.; Karplus, M. Collective Motions in Proteins: A Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations. *Proteins Struct. Funct. Bioinform.* **1991**, *11*, 205–217. [[CrossRef](#)]
60. Liu, W.S.; Wang, R.R.; Sun, Y.Z.; Li, W.Y.; Li, H.L.; Liu, C.L.; Ma, Y.; Wang, R.L. Exploring the Effect of Inhibitor AKB-9778 on VE-PTP by Molecular Docking and Molecular Dynamics Simulation. *J. Cell. Biochem.* **2019**, *120*, 17015–17029. [[CrossRef](#)] [[PubMed](#)]
61. Sun, Y.Z.; Chen, X.B.; Wang, R.R.; Li, W.Y.; Ma, Y. Exploring the Effect of N308D Mutation on Protein Tyrosine Phosphatase-2 Cause Gain-of-Function Activity by a Molecular Dynamics Study. *J. Cell. Biochem.* **2019**, *120*, 5949–5961. [[CrossRef](#)]
62. Selvaraj, C.; Omer, A.; Singh, P.; Singh, S. Molecular Insights of Protein Contour Recognition with Ligand Pharmacophoric Sites through Combinatorial Library Design and MD Simulation in Validating HTLV-1 PR Inhibitors. *Mol. Biosyst.* **2015**, *11*, 178–189. [[CrossRef](#)] [[PubMed](#)]
63. Satagopan, S.; North, J.A.; Arbing, M.A.; Varaljay, V.A.; Haines, S.N.; Wildenthal, J.A.; Byerly, K.M.; Shin, A.; Tabita, F.R. Structural Perturbations of Rhodospseudomonas Palustris Form II RuBisCO Mutant Enzymes That Affect CO₂ Fixation. *Biochemistry* **2019**, *58*, 3880–3892. [[CrossRef](#)]
64. Ashida, H.; Saito, Y.; Kojima, C.; Kobayashi, K.; Ogasawara, N.; Yokota, A. A Functional Link between RuBisCO-like Protein of Bacillus and Photosynthetic RuBisCO. *Science* **2003**, *302*, 86–290. [[CrossRef](#)] [[PubMed](#)]

65. Iñiguez, C.; Capó-Bauçà, S.; Niinemets, Ü.; Stoll, H.; Aguiló-Nicolau, P.; Galmés, J. Evolutionary Trends in RuBisCO Kinetics and Their Co-Evolution with CO₂ Concentrating Mechanisms. *Plant J.* **2020**, *101*, 897–918. [[CrossRef](#)]
66. Poudel, S.; Pike, D.H.; Raanan, H.; Mancini, J.A.; Nanda, V.; Rickaby, R.E.M.; Falkowski, P.G. Biophysical Analysis of the Structural Evolution of Substrate Specificity in RuBisCO. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30451–30457. [[CrossRef](#)] [[PubMed](#)]
67. Saito, Y.; Ashida, H.; Sakiyama, T.; de Marsac, N.T.; Danchin, A.; Sekowska, A.; Yokota, A. Structural and Functional Similarities between a Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (RuBisCO)-like Protein from *Bacillus Subtilis* and Photosynthetic RuBisCO. *J. Biol. Chem.* **2009**, *284*, 13256–13264. [[CrossRef](#)] [[PubMed](#)]
68. Ashida, H.; Saito, Y.; Nakano, T.; Tandeau De Marsac, N.; Sekowska, A.; Danchin, A.; Yokota, A. RuBisCO-like Proteins as the Enolase Enzyme in the Methionine Salvage Pathway: Functional and Evolutionary Relationships between RuBisCO-like Proteins and Photosynthetic RuBisCO. *J. Exp. Botany* **2008**, *59*, 1543–1554. [[CrossRef](#)]
69. Sugawara, H.; Yamamoto, H.; Shibata, N.; Inoue, T.; Okada, S.; Miyake, C.; Yokota, A.; Yasushi, K. Crystal Structure of Carboxylase Reaction-Oriented Ribulose 1,5- Bisphosphate Carboxylase/Oxygenase from a Thermophilic Red Alga, *Galdieria Partita*. *J. Biol. Chem.* **1999**, *274*, 15655–15661. [[CrossRef](#)]
70. Wang, Y.; Stessman, D.; Spalding, M. The CO₂ Concentrating Mechanism and Photosynthetic Carbon Assimilation in Limiting CO₂: How *Chlamydomonas* Works against the Gradient. *Plant J.* **2015**, *82*, 429–448. [[CrossRef](#)]
71. Banda, D.M.; Pereira, J.H.; Liu, A.K.; Orr, D.J.; Hammel, M.; He, C.; Parry, M.A.J.; Carmo-Silva, E.; Adams, P.D.; Banfield, J.F.; et al. Novel Bacterial Clade Reveals Origin of Form I Rubisco. *Nat. Plants* **2020**, *6*, 1158–1166. [[CrossRef](#)]
72. Alonso, H.; Blayney, M.J.; Beck, J.L.; Whitney, S.M. Substrate-Induced Assembly of *Methanococcoides Burtonii* D-Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Dimers into Decamers. *J. Biol. Chem.* **2009**, *284*, 33876–33882. [[CrossRef](#)]
73. Gunn, L.H.; Valegard, K.; Andersson, I. A Unique Structural Domain in *Methanococcoides Burtonii* Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco) Acts as a Small Subunit Mimic. *J. Biol. Chem.* **2017**, *292*, 6838–6850. [[CrossRef](#)]
74. Schreuder, H.A.; Knight, S.; Curmi, P.M.G.; Andersson, I.; Cascio, D.; Branden, C.I.; Eisenberg, D. Formation of the Active Site of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase by a Disorder-Order Transition from the Unactivated to the Activated Form. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9968–9972. [[CrossRef](#)] [[PubMed](#)]
75. Seno, Y.; Go, N. Deoxyoglobin Studied by the Conformational Normal Mode Analysis. I. Dynamics of Globin and the Heme-Globin Interaction. *J. Mol. Biol.* **1990**, *216*, 95–109. [[CrossRef](#)]
76. Levitt, M.; Sander, C.; Stern, P.S. Protein Normal-Mode Dynamics: Trypsin Inhibitor, Crambin, Ribonuclease and Lysozyme. *J. Mol. Biol.* **1985**, *181*, 423–447. [[CrossRef](#)]
77. Schloss, J.V. Comparative Affinities of the Epimeric Reaction-Intermediate Analogs 2- and 4-Carboxy-D-Arabinitol 1,5-Bisphosphate for Spinach Ribulose 1,5-Bisphosphate Carboxylase. *J. Biol. Chem.* **1988**, *263*, 4145–4150. [[CrossRef](#)]
78. Ng, J.; Guo, Z.; Mueller-Cajar, O. Rubisco Activase Requires Residues in the Large Subunit N Terminus to Remodel Inhibited Plant Rubisco. *J. Biol. Chem.* **2020**, *295*, 16427–16435. [[CrossRef](#)]
79. Watanabe, H.; Enomoto, T.; Tanaka, S. Ab Initio Study of Molecular Interactions in Higher Plant and *Galdieria Partita* Rubiscos with the Fragment Molecular Orbital Method. *Biochem. Biophys. Res. Commun.* **2007**, *361*, 367–372. [[CrossRef](#)]
80. Satagopan, S.; Chan, S.; Perry, L.J.; Tabita, F.R. Structure-Function Studies with the Unique Hexameric Form II Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase (Rubisco) from *Rhodospseudomonas Palustris*. *J. Biol. Chem.* **2014**, *289*, 21433–21450. [[CrossRef](#)] [[PubMed](#)]
81. Mueller-Cajar, O. The Diverse AAA+ Machines That Repair Inhibited Rubisco Active Sites. *Front. Mol. Biosci.* **2017**, *4*, 31. [[CrossRef](#)]
82. Taylor, T.; Andersson, I. Structural Transitions during Activation and Ligand Binding in Hexadecameric Rubisco Inferred from the Crystal Structure of the Activated Unliganded Spinach Enzyme. *Nat. Struct. Biol.* **1996**, *3*, 95–101. [[CrossRef](#)]
83. Atomi, H.; Fukui, T.; Kanai, T.; Morikawa, M.; Imanaka, T. Description of *Thermococcus Kodakaraensis* Sp. Nov., a Well Studied Hyperthermophilic Archaeon Previously Reported as *Pyrococcus* Sp. KOD1. *Archaea* **2004**, *1*, 263–267. [[CrossRef](#)]
84. Anwar, M.; Choi, S. Structure-Activity Relationship in TLR4 Mutations: Atomistic Molecular Dynamics Simulations and Residue Interaction Network Analysis. *Sci. Rep.* **2017**, *7*, 1–14. [[CrossRef](#)]