



HAL
open science

Structural Design and Analysis of the RHOA-ARHGEF1 Binding Mode: Challenges and Applications for Protein-Protein Interface Prediction

Ennys Gheyouché, Matthias Bagueneau, Gervaise Loirand, Bernard Offmann,
Stéphane Téletchéa

► To cite this version:

Ennys Gheyouché, Matthias Bagueneau, Gervaise Loirand, Bernard Offmann, Stéphane Téletchéa. Structural Design and Analysis of the RHOA-ARHGEF1 Binding Mode: Challenges and Applications for Protein-Protein Interface Prediction. *Frontiers in Molecular Biosciences*, 2021, 8, <10.3389/fmolb.2021.643728>. <hal-04064225>

HAL Id: hal-04064225

<https://hal.science/hal-04064225v1>

Submitted on 11 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Structural Design and Analysis of the RHOA-ARHGEF1 Binding Mode: Challenges and Applications for Protein-Protein Interface Prediction

Ennys Gheyouché¹, Matthias Bagueneau¹, Gervaise Loirand², Bernard Offmann¹ and Stéphane Téletchéa^{1*}

¹ UFIP, Université de Nantes, UMR CNRS 6286, Nantes, France, ² Université de Nantes, CHU Nantes, CNRS, Inserm, L'institut Du Thorax, Nantes, France

OPEN ACCESS

Edited by:

Joseph Rebehmed,
Lebanese American University,
Lebanon

Reviewed by:

Sophie Sacquin-Mora,
UPR9080 Laboratoire de Biochimie
Théorique (LBT), France
Guillaume Postic,
INSERM U973 Molécules
Thérapeutiques in Silico (MTI), France

*Correspondence:

Stéphane Téletchéa
stephane.teletchea@univ-nantes.fr

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 December 2020

Accepted: 13 April 2021

Published: 24 May 2021

Citation:

Gheyouché E, Bagueneau M,
Loirand G, Offmann B and
Téletchéa S (2021) Structural Design
and Analysis of the RHOA-ARHGEF1
Binding Mode: Challenges and
Applications for Protein-Protein
Interface Prediction.
Front. Mol. Biosci. 8:643728.
doi: 10.3389/fmolb.2021.643728

The interaction between two proteins may involve local movements, such as small side-chains re-positioning or more global allosteric movements, such as domain rearrangement. We studied how one can build a precise and detailed protein-protein interface using existing protein-protein docking methods, and how it can be possible to enhance the initial structures using molecular dynamics simulations and data-driven human inspection. We present how this strategy was applied to the modeling of RHOA-ARHGEF1 interaction using similar complexes of RHOA bound to other members of the Rho guanine nucleotide exchange factor family for comparative assessment. In parallel, a more crude approach based on structural superimposition and molecular replacement was also assessed. Both models were then successfully refined using molecular dynamics simulations leading to protein structures where the major data from scientific literature could be recovered. We expect that the detailed strategy used in this work will prove useful for other protein-protein interface design. The RHOA-ARHGEF1 interface modeled here will be extremely useful for the design of inhibitors targeting this protein-protein interaction (PPI).

Keywords: PPI, protein-protein docking, molecular dynamics simulation, ARHGEF1, RHOA

1. INTRODUCTION

Precise interactions between proteins allow a tight control on many functions and pathways, eventually leading to gene expression or silencing, to protein release or degradation, and even to cell death. To date, there are between 130,000 and up to 650,000 protein-protein interactions (PPIs) described (Ottmann, 2015), but only a fraction of PPIs is validated experimentally, ranging from 14,000 (Rolland et al., 2014) to 125,000 PPIs (<http://interactome3d.irbbarcelona.org/>, Mosca et al., 2013). This structural gap comes from the difficulties of obtaining experimentally full-length interacting proteins and then to resolving their structures using crystallography, NMR, or electron microscopy (EM). As a result, in most cases for a specific PPI, one has to combine existing incomplete experimental structures with *in silico* approaches. This virtual step is even critical for drug discovery, as exemplified with the successful targeting of 50 PPIs by small molecules (Skwarczynska and Ottmann, 2015). When no protein-protein structure is available, one has thus to perform protein-protein docking predictions.

The binding mode prediction of two proteins is very challenging since: (i) minor to major local structural rearrangements may be triggered upon protein recognition, (ii) one protein may recognize multiple proteins, and (iii) cofactors/nucleic acids may be involved to enhance or stabilize the interaction. The analysis of the various existing protein-protein interfaces available in the Protein Data Bank indicates also that (i) the interface area varies greatly between protein families, (ii) the composition in amino acids in this interface may be biased, and (iii) the binding lifetime is transient. Due to the complexity of modeling these diverse PPIs, there are many methods developed, and a global evaluation called CAPRI is performed periodically. We use the most robust and successful methods validated in this competition for protein-protein docking evaluation of the optimal binding mode of our example (Lensink et al., 2007, 2018, 2019).

To illustrate the process of modeling a protein-protein interface, we selected a protein complex where some unbound and bound experimental structures are available, and a complex between other members of the family is available. The first protein partner in our study is RHOA (gene *RHOA*), the second protein partner is Rho Guanine nucleotide Exchange Factor 1 (gene *ARHGEF1*). RHOA is a member of the RAS superfamily of small GTPases recognized as a master regulator of the actin cytoskeleton, thus driving multiple cellular processes, such as cell contraction, migration, proliferation, and gene transcription. While basal and controlled RHOA activity is required for homeostatic functions in physiological conditions, its uncontrolled overactivation plays a causative role in the pathogenesis of several diseases, such as cancer, neurodegenerative, or cardiovascular diseases (Guilluy, 2010; Cherfils and Zeghouf, 2011; Vetter, 2014; Loirand, 2015; Prieto-Dominguez et al., 2019; Arrazola Sastre et al., 2020). RHOA is a molecular switch that couples cell surface receptors to intracellular effector pathways by cycling between a cytosolic inactive state bound to guanosine 5'-diphosphate (GDP), and an active GTP-bound state that translocates to the membrane. The activation of RHOA is mediated by Rho nucleotide guanine exchange factors (GEFs) that promote the exchange of GDP for GTP, which are themselves turned on by the activation of upstream membrane receptors (Cherfils and Zeghouf, 2013). ARHGEF1 is the Rho GEF responsible for the activation of RHOA by angiotensin II through type 1 angiotensin II receptor in vascular smooth muscle cells (Loirand and Pacaud, 2010; Luigia et al., 2015). This signaling pathway participates in the physiological control of the vascular tone and blood pressure, and is causally involved in the pathophysiology of hypertension (Guilluy, 2010).

Small GTPases structure consists of a six-stranded β -sheet (β strands B1 to B6) linked by helices and loops (Ihara et al., 1998). In RHOA, the β -sheet is made up of the anti-parallel association of B1 and B2 and the parallel association of B3, B1, B4, B5, and B6, and there are five α helices (A1, A3, A3', A4, and A5) and three 3_{10} helices (H1–H3). RHOA possess two hinge regions, a

loop called switch I (29–42) and an helix called switch II (62–68), which are described to be more flexible than the core β -sheet (Dvorsky and Ahmadian, 2004), as shown in **Figure 1** for various bound and unbound experimental structures. RHOGEF proteins catalyze the exchange of GDP with 5'-triphosphate (GTP) on RHOA (Felline et al., 2019). Two domains on these proteins, Pleckstrin Homology (PH) and Dbl Homology (DH), are involved in the nucleotide exchange mechanism, the RHOA-bound DH domain being more rigid than the PH domain (**Figure 1C**). The DH domain consists of six α -helices arranged in an oblong shape, which interact with switch I and switch II regions of RHOA. The PH domain contains seven antiparallel β -strands forming a roll architecture, connected to helix $\alpha 6$ of the DH domain. The mechanism of nucleotide exchange involves large displacements of the PH domain relative to the DH-RHOA interaction (Felline et al., 2019).

In this article, we show how to combine virtual approaches with experimental data to predict reliably the formerly non-existing structure of a PPI. By using a rough structural superimposition or more advanced protein-protein docking methods, we build, analyze, and refine the RHOA-ARHGEF1 model and discuss the strengths and weaknesses of this approach. We, then, derive general recommendations to reproduce our approach to model other PPIs.

2. MATERIALS AND METHODS

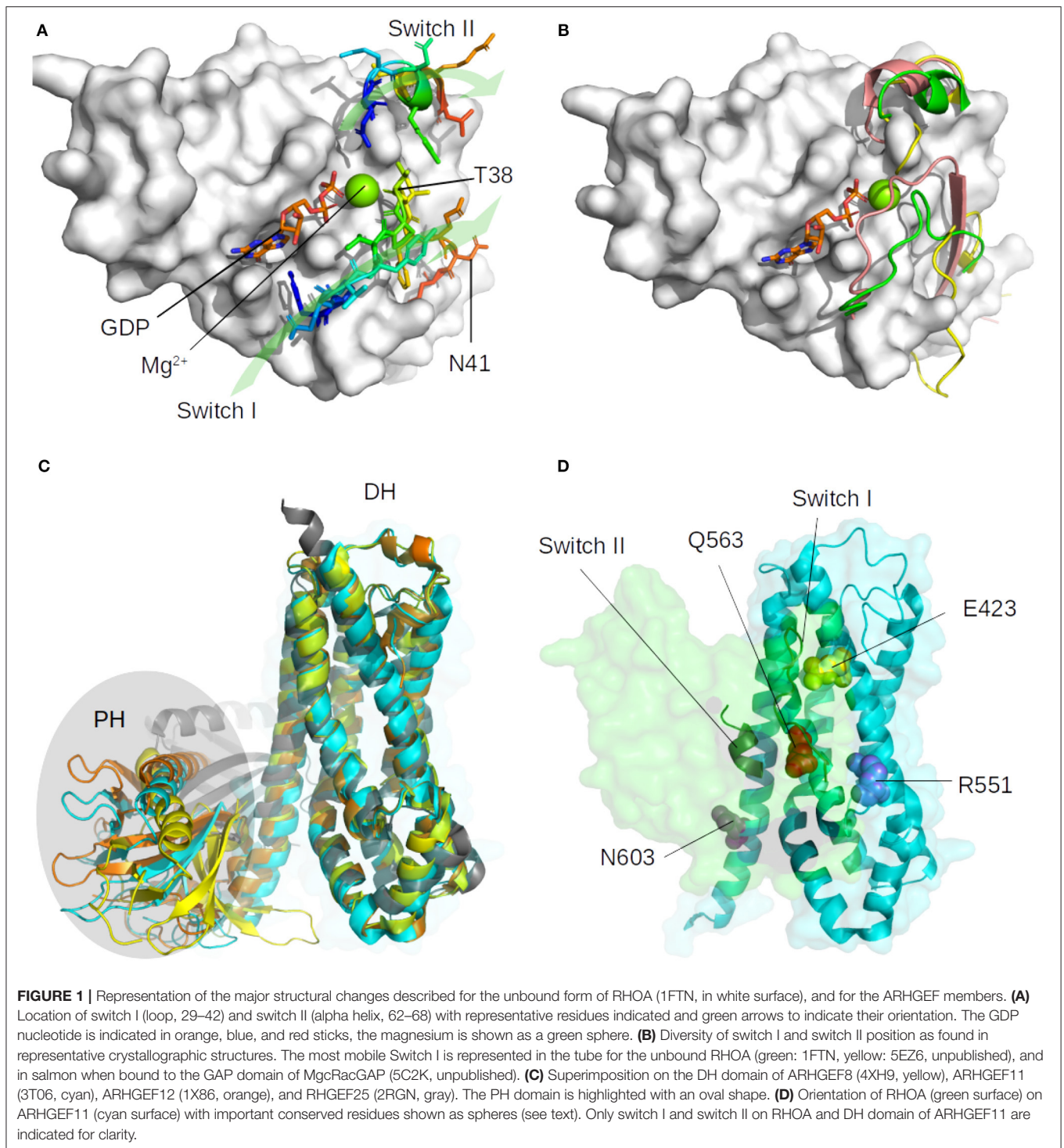
2.1. Sequence Analysis

A BLAST sequence search on the Non-Redundant (NR) database with RHOA or ARHGEF1 sequence as query was performed in June, 2018. The resulting sequences were aligned using clustal Omega (Madeira et al., 2019), amino acids conservation was estimated using Jalview (Waterhouse et al., 2009) and in-house scripts in Biopython (Cock et al., 2009). The phylogenetic analysis of human conserved sequences was performed in Genious version 2019.0.4 (<http://www.geneious.com/>).

2.2. Structure and Interface Analysis

We used all structures of RHOA complexed with all ARHGEFs available in the Protein Databank (Berman et al., 2003) in January 2018 and their respective unbound form when available. Only one representative chain by crystallographic structure was taken as reference (**Supplementary Table 1**). Experimental structures were analyzed using PDBePISA version 1.54 (Krissinel and Henrick, 2007; Krissinel, 2010). Two methods were selected to analyze protein-protein interfaces in order to obtain useful insights of the important residues involved in the interaction. The first one was 2P2I Inspector, version 2.0 (http://2p2idb.cnrs-mrs.fr/2p2i_inspector.html) (Basse et al., 2016) which computes a series of 51 chemical and physical descriptors from three-dimensional (3D) structures. The second one, PPCheck (<http://caps.ncbs.res.in/ppcheck/>) is a webserver for quantifying the strength of a protein-protein interface (Sukhwai and Sowdhamini, 2013). It can also be used to predict hotspots, perform computational alanine scanning, and to differentiate possible native-like conformations from the non-native ones

Abbreviations: RMSD, root mean square deviation; SAS, solvent accessible surface.



given a set of decoy ensembles as obtained through the protein-protein docking as it computes the strength of non-bonded interactions between any two proteins/chains present in the complex. Robetta Server was used to perform virtual alanine scanning of the interface (Kortemme et al., 2004). Models from docking or molecular dynamics simulations were visually

assessed and analyzed in The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. Root Mean Square Deviation (RMSD) was computed using the PyMOL `rms_cur` command. As RMSD is a global measure, we use two specific measures for rigid body docking and molecular dynamics simulations interface analysis as described in Takemura and

Kitao (2019): (i) Ligand-RMSD (L-RMSD) where one protein is Fixed (F) and the second protein is Mobile (M) to compute the L-RMSD. First the Fixed protein is superimposed on the same structure in the crystallographic reference, then the RMSD is computed on the Mobile (M) protein alone. (ii) Interface-RMSD (i-RMSD): in this case, only the amino acids known to be involved in the interface between both proteins are evaluated. Again, a first step consists of superimposing the one protein (RHOA or the PH domain of ARHGEFs) on the reference crystallographic structure to remove translation and rotation degrees of freedom potentially coming from the docking methods process. The angle between helices is computed using a plugin by Thomas Holder, which computes the angle between two vectors created from the coordinates of the C α atoms of each helix.

2.3. Superimposition Model of RHOA-ARHGEF1

A preliminary structure of RHOA bound to ARHGEF1 was derived from the co-crystal of RHOA-ARHGEF11 (PDB id:3T06) (Bielnicki et al., 2011). The complexes were created by superimposition of ARHGEF1 (3ODO) (Chen et al., 2011) on ARHGEF11 in PyMOL. This superimposition was submitted to MolProbity (Williams et al., 2018) for analysis and steric clashes were removed using Chiron (Ramachandran et al., 2011). Chiron performs rapid energy minimization of protein molecules using discrete molecular dynamics with an all-atom representation for each residue in the protein, this process allows to remove most of the steric clashes.

2.4. Protein-Protein Docking

The binding mode prediction was done in default mode for all methods using their respective webservers: (i) ATTRACT ff2g (<http://chemosimserver.unice.fr/attract/>) (Chéron et al., 2017), (ii) ClusPro version 2 (<https://cluspro.bu.edu/home.php>) (Kozakov et al., 2017), (iii) Haddock (van Zundert et al., 2016), (iv) PyDockWeb, Oct 2017 (<https://life.bsc.es/servlet/pydock/home/>) (Jiménez-García et al., 2013), (v) ZDOCK version 3.0.2f (Pierce et al., 2014).

2.5. Molecular Dynamics Simulation

All-atom simulations of unbound and bound proteins were performed using GROMACS 2016.3 (Abraham et al., 2015), the starting structures are detailed in **Supplementary Table 1**. Each system was prepared with the AMBER forcefield FF99SB-ILDN (Lindorff-Larsen et al., 2010) in explicit solvent (TIP3P) (Jorgensen et al., 1983) with a specific attention to protonation states as reported in PROPKA. A NVT followed by the anisotropic pressure coupling (NPT ensemble) protocol was applied until equilibration was reached and the full molecular dynamics simulation was computed for 500 ns up to 1 μ s. All simulations were run on the CCIPL cluster facility at the University of Nantes using GPUs. The force field parameters for GDP and GTP were gathered from the AMBER parameters database (<http://research.bmh.manchester.ac.uk/bryce/amber>) and converted to GROMACS format files using acpype (da Silva and Vranken, 2012). The resulting trajectories were visualized in the VMD version 1.9.1 (Humphrey et al., 1996), and GROMACS tools were used for various measurements.

TABLE 1 | Description of the protein-protein docking methods evaluated.

ATTRACT	<i>Ab-initio</i> protocol using a coarse-grained forcefield (ff2g) and manage to predict an estimation of the binding energy between the two proteins.	Chéron et al., 2017
CLUSPRO	FFT based program PIPER (pairwise potential based on the decoy at the reference approach) for the docking and using RMSD based clustering for filtering models then scoring using four different Scoring functions	Kozakov et al., 2017
HADDOCK	High Ambiguity-Driven DOCKing allows the use of external data, either experimental or from bioinformatics analysis to drive the modeling process	van Zundert et al., 2016
PyDockWeb	Rigid-body docking using FTDock, Gabb et al. (1997) and Jiménez-García et al. (2013) then an energy scoring based on empirical potential composed by of electrostatic and desolvation terms.	Cheng et al., 2007
ZDOCK	Fast Fourier Transform based protein docking program, version 3.0.2f (IFACE Statistical Potential, Shape Complementarity, and Electrostatics)	Pierce et al., 2014

PyContact was used to analyze protein contacts type, strength, and lifetime throughout the simulations (Scheurer et al., 2018). These contacts were plot with the R package MDplots (Margreitter and Oostenbrink, 2017).

3. RESULTS

In order to determine which docking method was the best for our specific needs, we have evaluated their performance (i) on recovering existing crystallographic structures of RHOA bound to a GEF, a process called re-docking, and (ii) on assembling the bound RHOA from a given crystallographic with a GEF from another crystallographic structure, this process being known as cross-docking.

3.1. Protein-Protein Docking Strategies

3.1.1. Z-Dock Is the Best Method for Building Our Complex

As docking strategies are based on different methods, it is difficult to determine *a priori* which method will produce the most reasonable starting complex for further studies. We assessed the top five performing docking software from CAPRI assessment, briefly introduced in **Table 1**, to produce RHOA/ARHGEF1 complexes: ATTRACT, ClusPro, HADDOCK, PyDockWeb, and ZDOCK.

3.1.2. Assessment of Webserver Performance in Re-docking Experiments

We evaluated the performance of each software according to its ability to recover the existing structure of bound RHOA and ARHGEFs. This method called re-docking allows to discriminate the accuracy of the algorithm studied on our system. The web servers define the first input protein given as the receptor, so

it stays fixed (F) and considers the second protein provided as the mobile protein (M). We performed our analysis with the small GTPase or the GEF as (F) or (M). The results are presented in **Table 2**. We computed the L-RMSD of the predicted complex by superimposing the Fixed protein with the same protein in the crystallographic structure and by computing the RMSD on the Mobile protein to assess the performance of each method. When comparing individually the predicted pose against the reference, ATTRACT and PyDockWeb perform equally with lower L-RMSD values for ATTRACT on its best predictions. ZDOCK and ClusPro present more diverse results and larger deviations. Although it should not be important, we observed that the input order of the fixed and mobile protein was affecting the prediction. This is of limited importance for ATTRACT and PyDockWeb, these methods are, therefore, less sensitive to the size of the mobile protein (RHOA being 200 AA long and GEF DH/PH domains being 600 AA long). We observe that for 4XH9, there is a dramatic decrease in the quality of the prediction although the calculations were performed three times. It is also possible to determine which models ranks the best among all poses and not only the first one: ZDOCK and ATTRACT proposes 10 binding modes, ClusPro offers multiple weights for their scoring functions for clusters of poses, and PyDockWeb presents the 100 best binding poses.

TABLE 2 | Analysis of docking software performance for complex formation using small GTPase (RHOA) or GEF (ARHGEF 8/11/12/25) as a mobile (M) or fixed (F) structure.

GEF(F). RHOA(M)/ GEF(M). RHOA(F)	ZDOCK	ATTRACT	ClusPro	PyDockWeb
3T06	1.60/1.92	0.70/0.97	3.21/7.10	1.66/1.34
4XH9	1.99/3.06	20.25/14.09**	4.27/1.80	1.52/1.53
1X86	2.61/3.23	1.11/1.71	3.32/5.81	1.44/0.92
2RGN	2.14/3.00	1.46/2.24	3.93/3.01	0.92/0.80
Computing time	24 h	24 h	24 h	1 week

The L-RMSD values are in (Å). The best prediction analyzed is the first of the best 10 poses or the best cluster of poses classified by each method. The best predictions are in bold, the wrong predictions are in italics. HADDOCK could not be evaluated at this stage since we chose to not use data guidance. Computing time is indicative of one docking experiment.

**Calculation were performed independently in triplicate to exclude any temporary issue.

All the poses identified as close to the crystallographic structure with a global RMSD under 2 Å were present in the top 10 solutions of the best cluster for each method. Altogether, this lead us to consider the PyDockWeb at the end of the re-docking study, although its computation time is significantly higher than the other methods (**Table 2**). As ZDOCK and PyDockWeb provided more reliably poses close to the original crystallographic structures, these two methods were kept for the following analysis.

3.1.3. Assessment of the Best Performing Methods in Cross-Docking Experiments

We verified the dependence on the initial structure for the protein-protein binding mode prediction. We applied a cross-docking experiment where each bound RHOA in one crystal structure is evaluated against another GEF partner. We used the free RHOA structure as a sensitivity control for our cross-docking measurements. There is no dependence for the docking result linked to the pdb input for RHOA or GEF. The results are available in **Table 3**: as one would expect, it is more complicated to build hybrid complexes than re-docking complexes. Out of the 12 combinations of partners for the cross-docking, ZDOCK is better for six predictions, PyDockWeb for five, and their prediction is good and close in one case. For two cases, pydockweb finds a very different orientation than the crystal structure : RHOA: 1X86/GEF: 2RGN (30.27 Å), RHOA: 4XH9/GEF: 1X86 (34.25 Å), where ZDOCK finds a close conformation (3.19/3.31). For both methods, the docking of the unbound RHOA produces very unrealistic protein-protein interfaces, indicating the sensitivity of the method toward switch I and II adaptations of RHOA for GEF binding although the RMSD between bound RHOA (3T06) and unbound RHOA (1FTN) is small (0.7 Å).

Considering the results of the re/cross-docking, we selected ZDOCK as the best method to predict the more favorable binding mode between RHOA and ARHGEF1 (**Figure 2**).

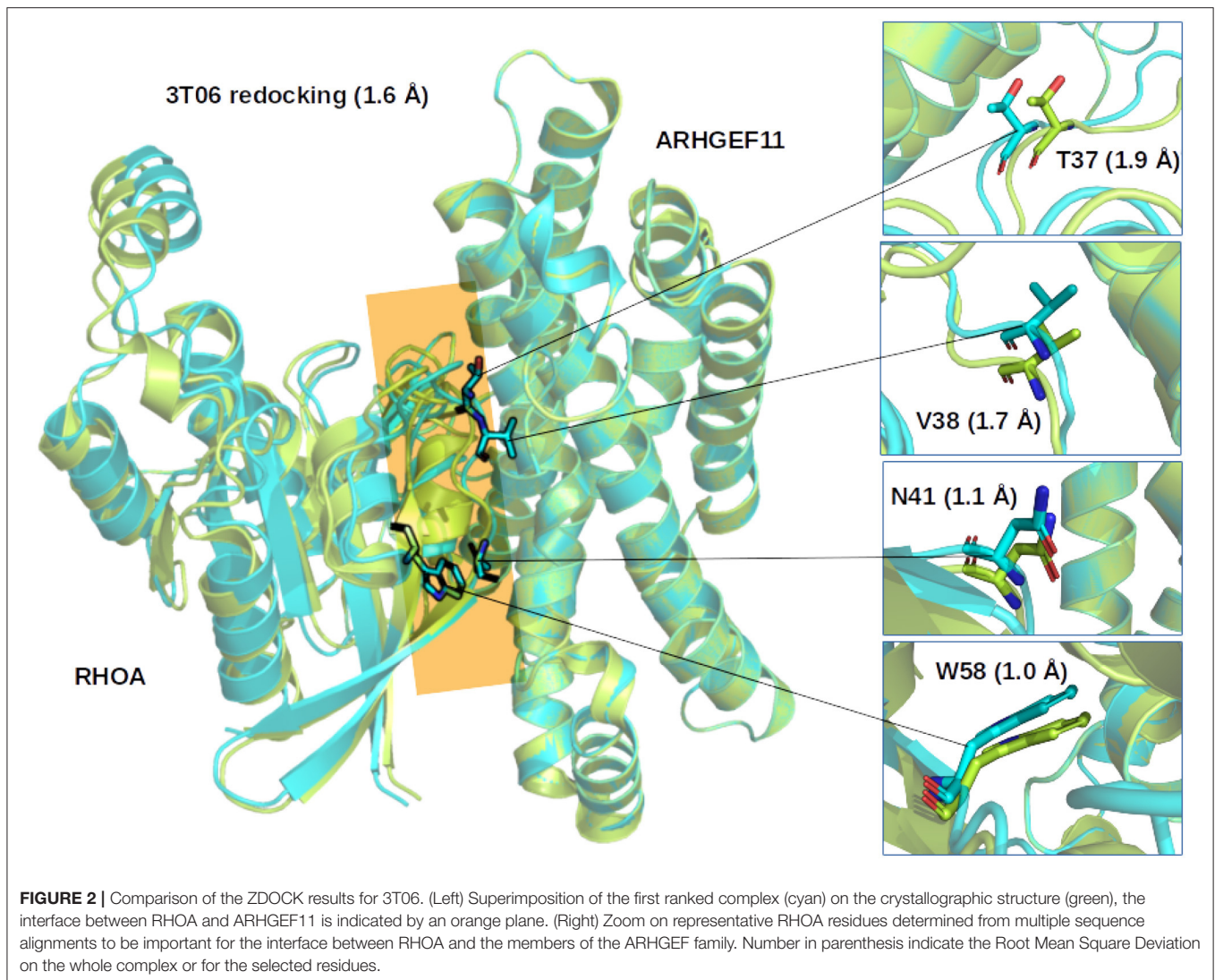
3.1.4. Evaluation of the Best Models Based on Known Interactions to Select the Best RHOA Candidate Structure

Since no experimental structure of the RHOA-ARHGEF1 interface is available, we used the crystallographic structures

TABLE 3 | Analysis of the crossdocking performance for the ZDOCK and PyDOCKweb. RHOA was considered as the ligand (M), and each GEF was fixed (F).

ZDOCK/pydockweb	GEF			
	GEF11 (3T06)	GEF8 (4XH9)	GEF12 (1X86)	GEF25 (2RGN)
Bound	3T06	<u>1.60/1.66</u>	2.15/2.10	3.22/ 1.86
RHOA	4XH9	3.35/5.58	<u>1.99/1.52</u>	3.31/34.25
	1X86	3.15/ 2.46	2.66/ 2.02	<u>2.61/1.44</u>
	2RGN	3.01/5.38	2.39/3.38	<u>2.53/2.16</u>
Free RHOA	1FTN	14.42/29.35	17.02/16.39	31.08/38.60
				14.44/38.13

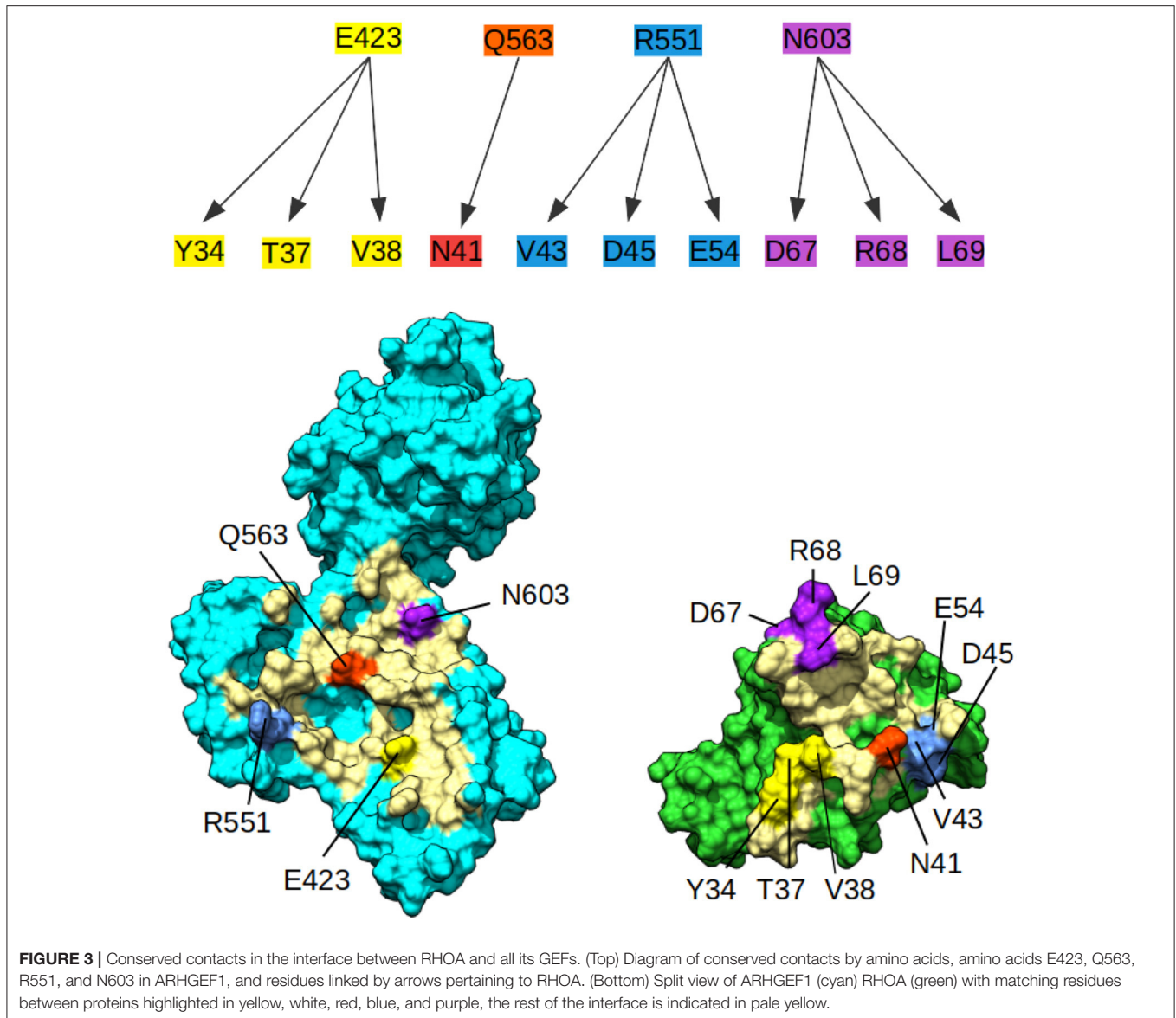
The RMSD value (in Å) for the mobile part is displayed for the best prediction in the first 10 poses. The results for re-docking experiments, underlined, are identical to **Table 2**. The best predictions are in bold.



of other GEF paralogs bound to RHOA to determine which amino acids are shared in all complexes. The amino acids mapping to RHOA and ARHGEF1 was done using multiple sequence alignments comparison (data not shown). Four residues are conserved in GEFs interfaces, namely, E423, Q563, R551, and N603 in ARHGEF1. As can be seen in **Figure 3**, these shared residues for all GEF interfaces can be split in zones or individual amino acids contacts. By analogy to existing complexes, ARHGEF1 E423 has to be present close to Y34/T37/V38 (a region called switch I in RHOA), R551 has to be close to V43/D45/E54 of RHOA, and N603 has to be close to D67/R68/L69 (a region called switch II in RHOA). Only one amino acid in RHOA, N41 seems to bind exclusively to Q563. It is well-established that RHOA is very rigid due to the strong structural requirements imposed by the GTP recognition and hydrolysis mechanism (Dvorsky and Ahmadian, 2004). Only two regions called switch I and switch II are more flexible with or without binding partners, as seen in **Figure 1**. Our study allows a more detailed understanding of

the interaction between amino acids pairs important for the RHOA-ARHGEF binding.

As done for ARHGEF1, we also analyzed the interface residues of the other GEFs present in each crystallographic structures in the complex with RHOA and assessed their amino acids conservation using multiple sequence alignments. Four residues of ARHGEF1 are present at the interface, and 10 for RHOA. As can be seen in **Table 4**, we have listed all amino acids present in interaction between both proteins, i.e., at least one amino acid of RHOA is in contact with one amino acid or more of a given ARHGEF. On average, the number of contact pairs recovered after docking represents at least half of the residues known to be present on both sides of the interface. This indicates that our docking strategy allows to build a reasonable starting structure for the RHOA-ARHGEF1 complex. As sequence conservation on the DH+PH domains modeled here is the most important among ARHGEF11, ARHGEF12, and ARHGEF1 (**Supplementary Table 2**), and provided the docking validation steps showed that ARHGEF11 docking



allowed to recover more contacts than with ARHGEF12 docking, we chose the RHOA structure found in 3T06 to dock it using ZDOCK with the unbound structure of ARHGEF1 (3ODO). The resulting complex will be referenced thereafter as complexD (Docking).

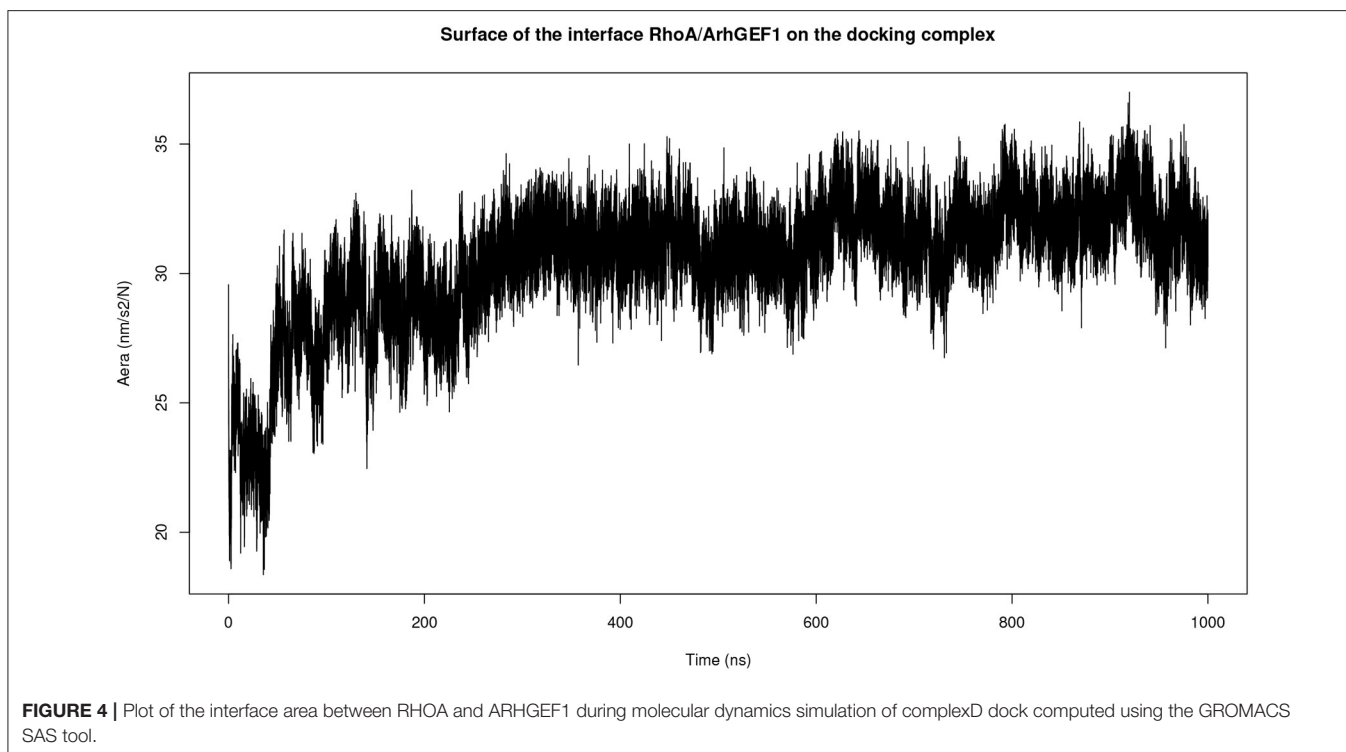
3.2. Template-Based Complex Modeling

Since there are some experimental structures of RHOA bound to ARHGEF1 homologs, we also predicted the bound structure of both proteins with a simpler approach, based on structural superimposition in PyMOL. We first used a rhoA-bound crystal structure and superimposed the free ARHGEF1 on all the homologous GEFs. By doing so with rigid models, we could not take into account the concerted induced-fit required for finely tuning the interaction. We used a webserver from Dokholyan Team named Chiron which allows to relax

the most important steric clashes (<http://redshift.med.unc.edu/chiron/>) (Ramachandran et al., 2011). In order to solve all the bumps, many rounds were necessary. A preliminary structure of RHOA bound to ARHGEF1 was derived from the RHOA-ARHGEF11 crystal structure (3T06). The complexes were created by superimposition of ARHGEF1 (3ODO) on ARHGEF11 in PyMOL. This superimposition was submitted to MolProbity for analysis, and steric clashes were removed using Chiron. From there, we used classical descriptors to evaluate the resulting complex (delta SAS, RMSD, ...) with a special look into the interface size. This interface was analyzed with PDBePISA. The best binding was found when using ARHGEF1 from 3ODO and RHOA from 3T06: the interface is 2,949 Å² corresponding to around 5% of the total surface of ARHGEF1 and to around 11% of RHOA total surface area. In the complexD, this interface comprises amino acids 3 to 181 from RHOA and 392 to 761

TABLE 4 | Numbers of amino acids determined to be in interaction between both proteins at the interface, numbers from RHOA (x/10) or its complexed ARHGEF (y/4), as found in the different binding mode generated by ZDOCK during the crossdocking experiments, while the ARHGEF partner was kept fixed.

	ARHGEF12 (1X86)	ARHGEF25 (2RGN)	ARHGEF11 (3T06)	ARHGEF8 (4XH9)
RHOA (1X86)	5/10-3/4	5/10-2/4	2/10-1/4	3/10-1/4
RHOA (2RGN)	3/10-3/4	6/10-3/4	8/10-4/4	5/10-2/4
RHOA (3T06)	5/10-1/4	7/10-3/4	6/10-3/4	6/10-2/4
RHOA (4XH9)	3/10-3/4	4/10-2/4	1/10-1/4	7/10-3/4
RHOA (1FTN)	0	0	0	0



from ARHGEF1, for a total interface size of 2,830 Å². Those values are smaller than the values found for the homologs complexes where this interface area is on average 3,371 Å². Before exploring further the complexD and complexT, we verified our models with PPcheck, which decomposes the interaction energy in three terms: (i) hydrogen bonding (Ehyd), (ii) inter-chain van der Waals interactions (Evw), and (iii) inter-chain electrostatic interactions (Eele). The total stabilizing energy is then divided by the total number of interface residues to obtain the energy per residue. No significant deviation requiring further refinements with the Chiron webserver was present in both models.

3.3. Molecular Dynamics Simulation Refinement for ComplexT and ComplexD

Both complexes were modeled using molecular dynamics simulations to determine if the interface could be refined during this procedure. The initial surface area in complexD, 2,830 Å² at the beginning of the simulation, stabilizes to

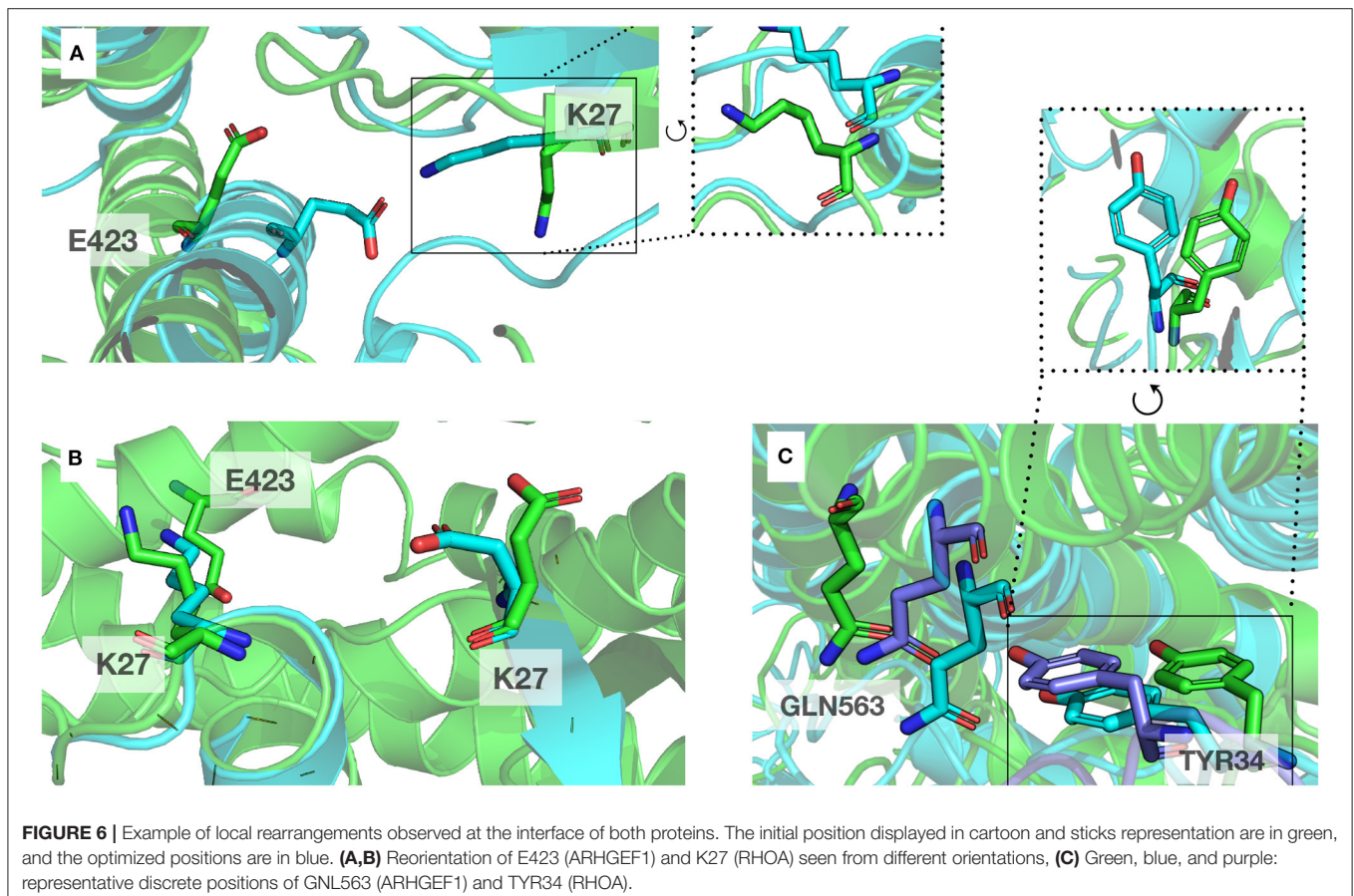
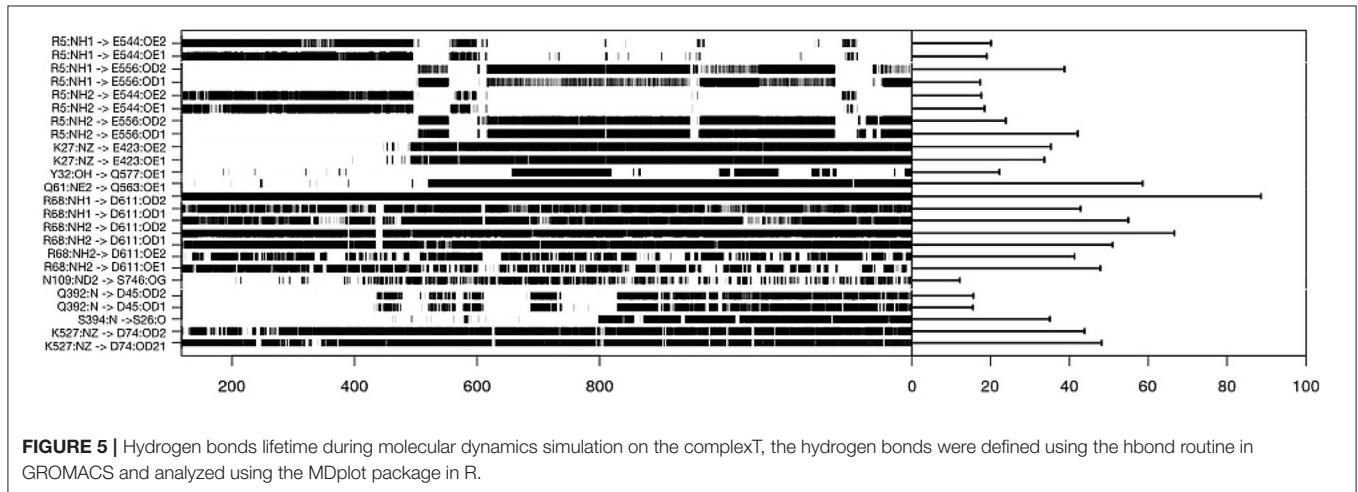
3,056 Å² between 200 and 1,000 ns of the simulation (**Figure 4**). The molecular dynamics simulation of RHOA-ARHGEF1 complexT also remained stable for most of the time with a rapid initial increase in the interface area followed by a plateau after 250 ns. The average interface size in this plateau is 3,150 Å² (data not shown). Starting with two different models, we observe an augmentation in protein surface contact driven by local adjustments. When considering individually each protein at the RMSD level, there is a higher deviation for RHOA than ARHGEF1, implying that RHOA undergoes most of the conformational changes, as we will see in details below.

3.4. Interface Contacts Evolution Over Time

To understand the evolution of interface complex during the simulations, we analyzed the hydrogen bonds between the two partners. The result is shown in **Figure 5** where we only plotted hydrogen bonds with a lifetime in the simulation

over 15%. With this plot, we can identify which amino acids, side chains are seeing important rotations. For instance, an important hydrogen bond is conserved between RHOA-ARG5 and ARHGEF1-GLU544 or ARHGEF1-ASP556. The most stable contact is RHOA-ARG68—ARHGEF1-ASN603/ASP611, since it is observed for 900 ns, or 90% of the simulation time. For others contacts, some were present from the beginning

of the simulation, others appeared and disappeared. Since it may take time to stabilize contacts, we observe that an important interaction appears between RHOA-GLN61 and ARHGEF1-GLN563 at 500 ns. Interestingly, some of these hydrogen bonds are members of the very conserved list of amino acids listed above, for instance, for RHOA-ARG68 and ARHGEF1-ASN603.



4. DISCUSSION

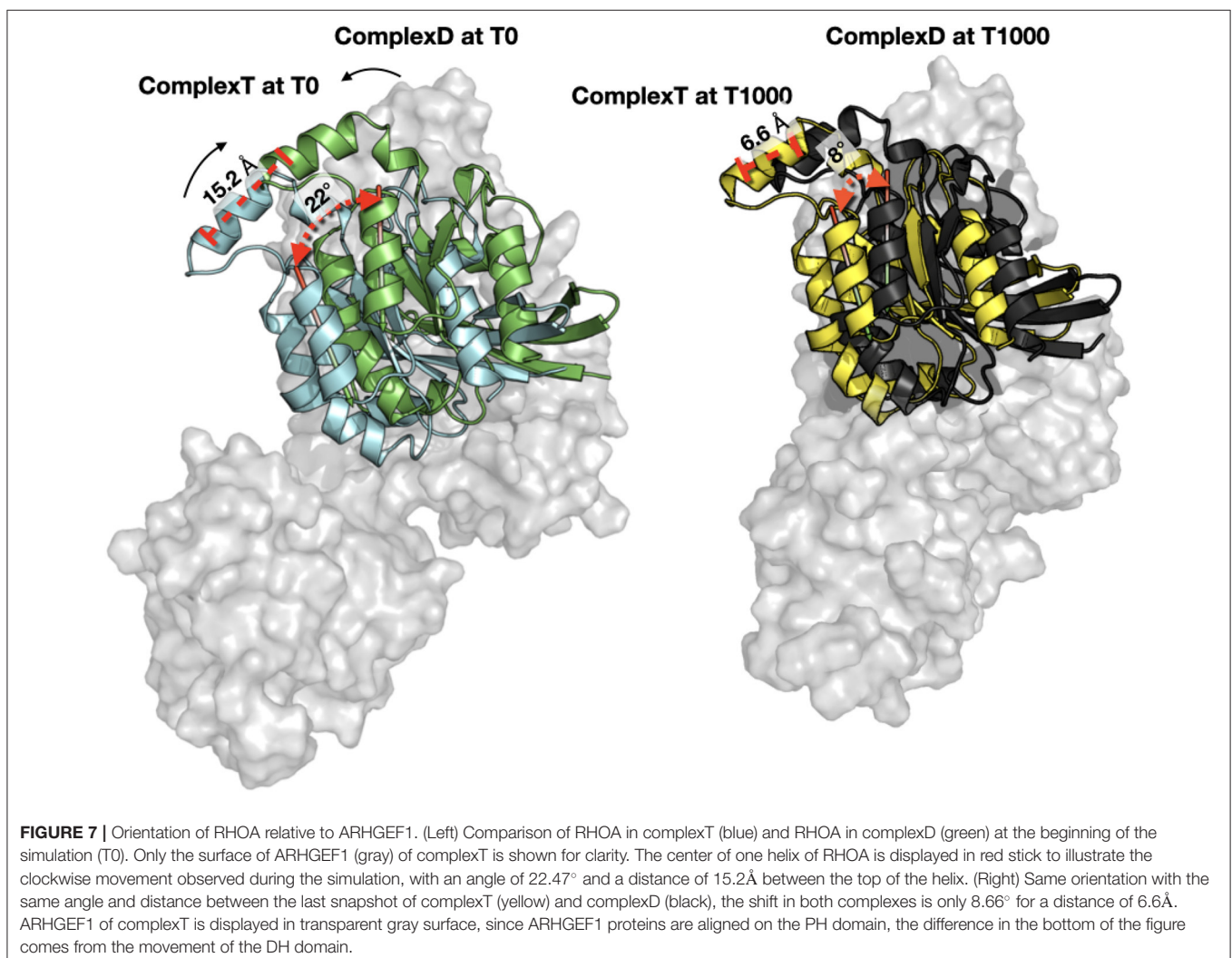
Since no experimental structure of RHOA-ARHGEF1 was available from X-ray studies, NMR, or EM studies, we had to model it. Using the protein sequences of Rho and GEF families, and the existing protein structures of bound homologs RHOA-ARHGEF8 (Petit et al., 2018), RHOA-ARHGEF11 (Bielnicki et al., 2011), RHOA-ARHGEF12 (Kristelly et al., 2004), and RHOA-ARHGEF25 (Lutz et al., 2007), we analyzed to determine which amino acids were shared at the interface of the complex (Figure 3). This sequence and structure-based information was important to assess the validity of our models.

4.1. Initial Models of RHOA-ARHGEF1 Complex

The prediction of protein-protein interface using docking methods is still an important field of research (Smith and Sternberg, 2002; Lensink et al., 2007) but the predictive power of these methods greatly varies depending on the protein families (Bendell et al., 2014; Wang et al., 2017). As no GEF

or RHOA experimental structure was used as target to assess the methods in recent CASP experiments, we benchmarked how these methods could perform on our specific case, using re-docking and cross-docking experiments. We selected ZDOCK after a careful quantification and inspection of the re-docking/cross-docking experiments since its results were the most robust across most predictions and in agreement with our sequence+structure derived data. The best model (**complexD**) selected from ZDOCK contains 5 out of 10 shared amino acids in RHOA and 3 out of 4 shared amino acids from ARHGEF1, with an interface surface area of 2,830 Å².

As the docking experiments are time-consuming and contain also uncertainties, we did also a more crude approach using PyMOL. We analyzed the existing structures of RHOA bound to other members of the ARHGEF family to find the best starting template for our structural comparison. A superimposition of ARHGEF1 (3ODO) on RHOA-ARHGEF11 (3T06) was then performed in PyMOL and further refined using Chiron (Ramachandran et al., 2011). This modeled interface



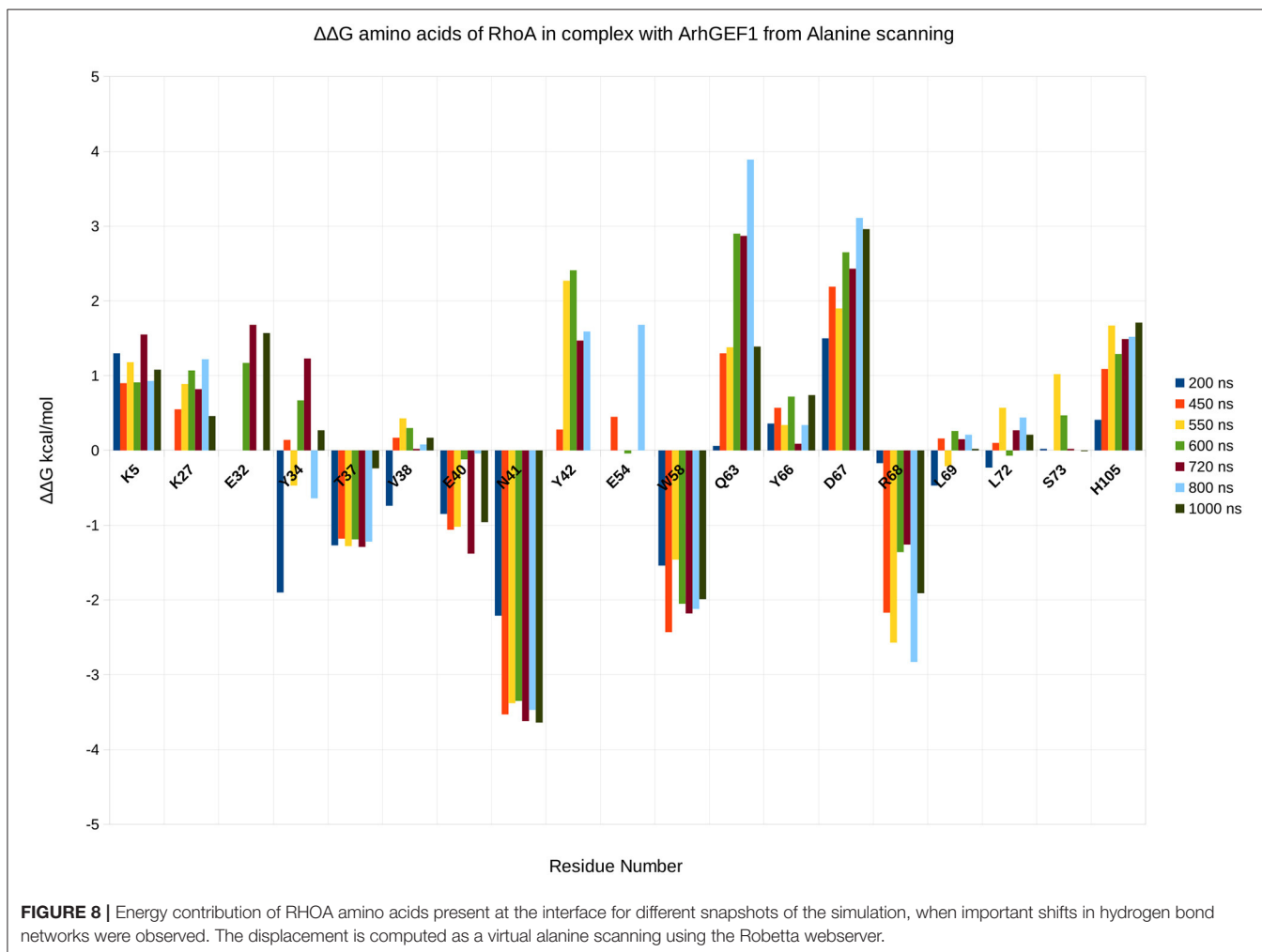
(**complexT**) allowed to correctly find the position of 2 out of 10 amino acids in RHOA and 2 out of 4 amino acids from ARHGEF1.

Both methods allowed to define comparable starting complexes of the RHOA-ARHGEF1 interface from rigid templates. Major steric clashes were carefully examined using Chiron (Ramachandran et al., 2011) and visual inspection, but no further amino acids adjustment was required. The RMSD difference between both models is 0.4 Å for ARHGEF1 and 9.5 Å for RHOA. This larger difference in RHOA position comes from an alternate orientation of the protein relative to ARHGEF1 with a clockwise rotation of 22° between RHOA in **complexT** and RHOA in **complexD** (**Figure 7**). This alternative positioning of RHOA in comparison to other members of the GEF family is also present in crystallographic structures. Both **complexT** and **complexD** seemed therefore reasonable starting complexes, with a comparable building time of 1 day for both protocols: instant for PyMOL superimposition plus 1 day for removal of clashes in Chiron and 1 day for ZDOCK prediction.

4.2. Molecular Dynamics Interface Refinement

A classical method to enhance protein models is molecular dynamics simulations (MD) (Mirjalili et al., 2014). We performed MD on **complexT** and **complexD** for 1μs each. During this simulation of the **complexT**, the interface area in the complex increased (3,480 Å²) in comparison to the initial complex (2,949 Å²). We identified amino acids conserved in all RHOA-ARHGEF complexes by combining structural sequence analysis (**Figure 3A**). These contacts are stable throughout the simulation (R5, R68/E544, D556, N603, and D611). Interestingly new contacts are observed K27-E423, R68-D611 led by local rearrangements of amino acids, in particular K27, Y34 (RHOA), and E423 (ARHGEF1) (**Figure 6**).

Both **complexT** and **complexD** lead to a similar RHOA-ARHGEF1 interface at the end of the simulation. At the beginning of the simulation ($t = 0$ ns), **ComplexD** and **complexT** only have a global RMSD difference of 1.4 Å between them, but at the end ($t = 1,000$ ns), the RMSD rises to 3.7 Å. Since the RMSD is a global measure of the movements, i.e., both proteins moved, it



is important to understand how proteins evolved independently during the simulations.

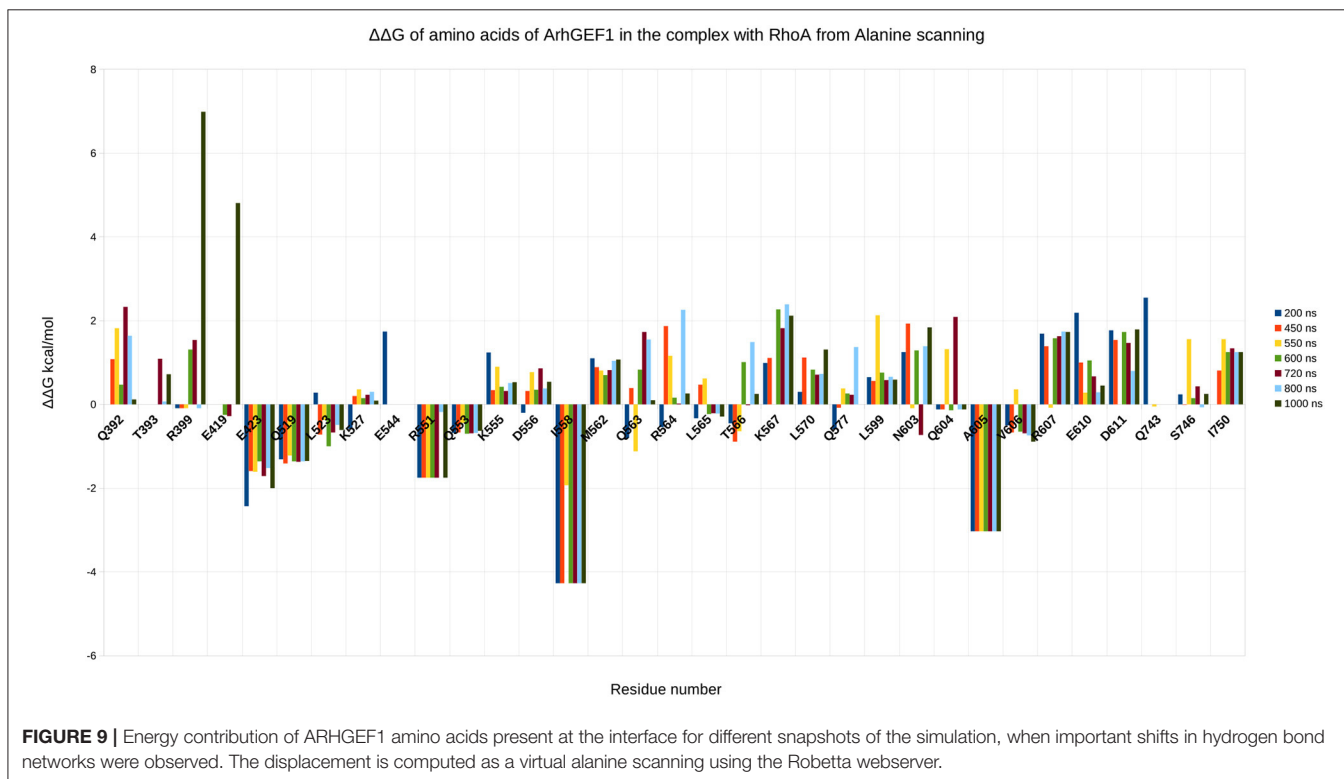
When each trajectory is taken individually, we observe that ComplexD moved more (5.1 Å) from its initial structure than complexT (2.9 Å). This apparent difference comes mostly from a larger movement of the PH domain in the complexD simulation since the RMSD between the initial structure and the end of the simulation concerning only the ARHGEF1 protein is 18 Å. This apparently large difference in ARHGEF1 position for complexD is the consequence of two movements: (i) the local rearrangements of the DH domain (the core RHOA binding domain of GEFs), which is similar in both complexD and complexT (6 Å), and (ii) a larger movement of the PH domain which may be involved in the nucleotide exchange. When aligning the DH domain in both trajectories, we, therefore, see a more important rotation of RHOA relative to ARHGEF1 for complexD (6 Å) than for complexT (3.5 Å) as illustrated in **Figure 7**. The main interface enhancements thus appear locally at the ARHGEF1-RHOA interface, mostly on the DH domain of ARHGEF1, and more globally with a clockwise (+22°) rotation in the complexT, and a slighter anti-clockwise (−3°) rotation in complexD.

Both complexes are refined after molecular dynamics simulations, many important amino acids saw an increase in contact frequency and the position of RHOA relative to ARHGEF1, either inherited from the superposition onto ARHGEF11 or from the docking studies with ZDOCK, led to a strong convergence of the interaction. This study confirms

the interest of using molecular dynamics simulation to increase model quality.

4.3. Validation of the Binding Mode

To validate the binding mode from the simulations, we used the interactions as a starting point and analyzed specific interaction of this complex found in the literature. We found back couples of interactions, which have a lifetime of over 15%, are conserved in all RHOA/GEFs binding modes, with some interactions specific to the RHOA-ARHGEF1 interface. We observed that the complex tends to go toward a more stable conformation when the PH domain moves to enclose RHOA, with an increase in the number of interactions, with a mean SAS going over 3,100 Å², the mean surface area for all complexes of RHOA/GEFs available so far. We could identify some specific contacts for RHOA-ARHGEF1 from the complexT, namely D59-K567, Q63-T566, which were not described previously (Hoffman and Cerione, 2002; Derewenda et al., 2004). In the complexD, the specific contacts E97-S746 has already been described by Gasmi-seabrook (Gasmi-Seabrook et al., 2010) as an essential contact in the nucleotide exchange for PDZ-ARHGEF1 and RHOA. This contact is not observed in the complexT simulation. Starting from two complexes built with different strategies, we were able to have a perfect compatibility between experimental predictions and our *in silico* methods. The identified additional contacts, specific for the RHOA-ARHGEF1, will require experimental exploration since there were differences from the two simulations, potentially coming for the overall dynamics of the interface. To guide experimental validations, we



studied *via* virtual alanine scanning if some amino acids could be qualified as hotspots (Kortemme et al., 2004; Jiang et al., 2017) of the interface (Figures 8, 9). Only three amino acids seem to contribute strongly to the binding of both proteins: (1) N41 for RHOA, already identified by multiple sequence alignments and structure comparison, (2,3) I558 and A605 in ARHGEF1. These residues seem to be robustly involved in the interface during all the simulations, either starting with the complexD or the complexT.

4.4. Selection of the Best Model

The knowledge acquired with our strategies helped us to understand the most relevant elements for the binding of the two partners altogether with insights for selecting/computing relatively good refined models. In the initial models after minimization for complexT, there were already 5 over the 10 conserved (E423, Q563, and N603) contacts and for the complexD 4 over 10 (mostly with E423). After simulation for complexT, we can see 8 over 10 conserved contacts and for complexD, there are 6 over 10 contacts during a short time frame where the interface SAS is the highest. Some contacts are seen only thanks to the simulations and one question rises, what is more important to select for qualifying the best model? Its higher number of contacts or the presence

of conserved/important contacts? In our case, both models show conserved/important contacts and new contacts specific to each model. The amino acids detected as hotspots are not conclusive since they are present in both trajectories. During the simulations, even if starting from somewhat different structures, the interaction between ARHGEF1 and RHOA converges. Our model building strategy clearly indicates that a molecular dynamics simulation, starting from rationally designed PPI, improves the initial models.

4.5. Comparison of the MD Model With Information-Driven HADDOCK Docking

HADDOCK is very efficient for protein-protein docking when experimental/bioinformatics constraints can be added for driving the docking. As we had determined the important residues for the binding interface, we used them in HADDOCK webserver first for redocking experiments on the 3T06 crystal structure as shown for other methods in Table 2, and obtained a RMSD of 0.75 Å with 90% of the structure in the first same cluster, better than all other protein-protein docking methods. We then did the prediction of the interaction model using ARHGEF1 from 3ODO and RHOA from 3T06. This prediction gave us nine clusters, and after careful analysis only two seemed to have the expected binding mode: cluster 1 and cluster 5. When

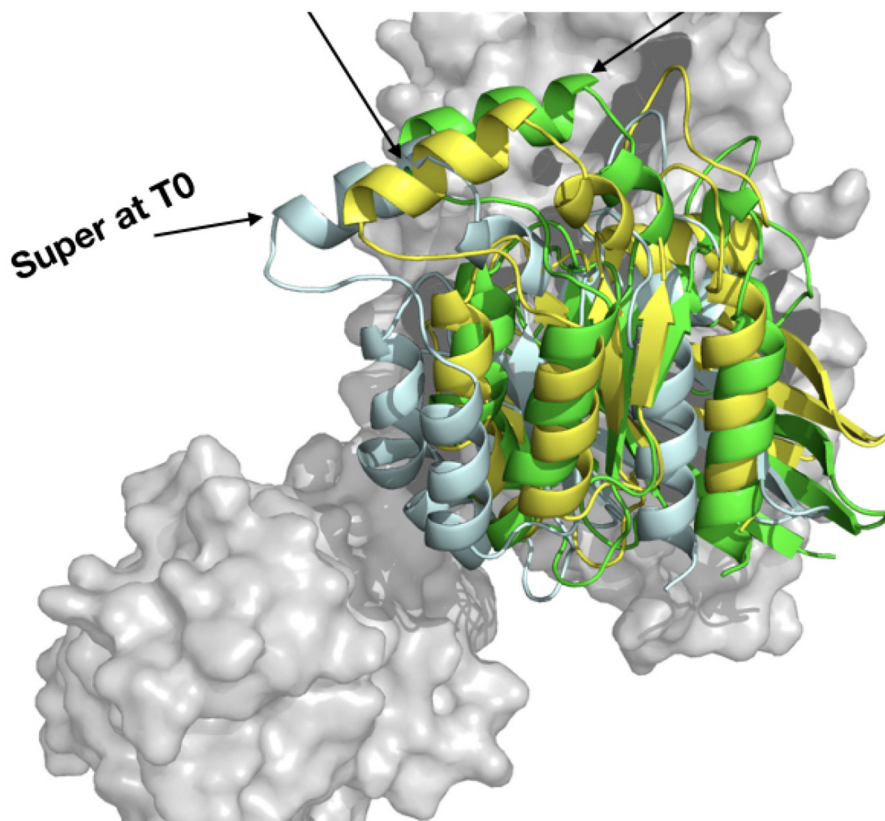


FIGURE 10 | Comparison of the HADDOCK webserver cluster 1 model built from 3T06 RHOA and 3ODO ARHGEF1 (yellow), complexT (light cyan), and complexD (green).

cluster 1 is compared to ZDOCK's derived complexD (without information driven construction), this cluster 1 has a RMSD of 0.62 Å, also very close to complexT with a RMSD value of 0.65 Å (Figure 10). Qualitatively, the cluster1 model displays 6 over 10 of the contacts given as input. If experimental data are available, for instance, coming from mutagenesis experiments, HADDOCK allows their incorporation to guide the binding mode. In this situation, HADDOCK is certainly the best strategy to build a PPI, provided these data can be transformed in sufficient constraints as input. However, the resulting model provided by HADDOCK in 1 day compared to other docking methods is very interesting. The interface area for cluster 1 is 1,478 Å², slightly better than complexD model before refinement (1,420 Å²), but far from the refined interface obtained after molecular dynamics simulations (>3,000 Å²). Even if it is possible to build a reliable model by integrating various data in HADDOCK, a long molecular dynamics simulation, with a simulation time above 250 ns is still required to enhance the quality of a PPI (Feig and Mirjalili, 2016).

5. CONCLUSION

Most biological processes involve transient protein-protein interactions, in particular for cellular signaling. The RHOA-ARHGEF1 interaction is responsible for the activation of RHOA downstream of type 1 angiotensin II receptor signaling in vascular smooth muscle cells, thereby controlling vascular tone and blood pressure (Loirand, 2015).

Our study aims at exemplifying how one can model a protein-protein interaction when sufficient experimental structures are present, but only experimental data for close homologs are available. We set up two different strategies summarized in Figure 11. One has first to identify the close homologs. If the members of the family have a standardized name, they should rapidly be identified directly in the Protein Data Bank (Berman et al., 2003). If not, a search on the National Center for Biotechnology Information (NCBI) structure service, in the Protein Families Database (Mistry et al., 2021), or in the PALI database (Balaji et al., 2001) should help in finding close homologs. If not, the protocols described in our work should be considered with caution. When the close homologs are identified, it is possible to apply the protocols previously described. The first one, based on structural superposition of partners, allows a rapid building of the complex, but provides required local adjustments to avoid steric clashes. We expect it to be useful for a preliminary study of how the proteins interact. The second strategy based on most advanced methods combining the search of the best binding mode *via* the assessment of the results of protein-protein docking, followed by the refinement of the best docked model using molecular dynamics simulations. This model showed not only increased shape complementarity and

increased contacts but also provides insights into the dynamics of the detailed amino acids interactions between the partners. This more advanced strategy is probably only accessible to experts and should only be required for atomic-level analysis and mechanistic studies. In our study, both strategies gave close initial models, but we do not expect the results on RHOA-ARHGEF1 to be amenable for general purpose. We, therefore, recommend to use a protein-protein rigid-body docking study (complexD) for producing the initial interaction mode. In our study, ZDOCK was better if precision, robustness, and time are taken altogether into consideration. When possible, we recommend to perform long molecular dynamics simulations to enhance the network of interaction between both proteins and to get a better overview of the lifetime of each interaction. More generally, we expect these strategies will be successfully applied to a variety of targets where a partial structural coverage of both partners is known, provided the complex to model has characteristics comparable with the two proteins described in this article.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

ST and GL obtained the funding. EG did the experiments and analysis. BO and GL supervised the work. EG and ST wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was part of the PIRAMID project funded by the French Regional Council of Pays-de-la-Loire and the TROPIC project, supported by ANR grant Programme d'Investissements d'Avenir (ANR-16-IDEX-0007), the French Regional Council of Pays-de-la-Loire, and Nantes Métropole.

ACKNOWLEDGMENTS

The authors thank the Center de Calcul Intensif des Pays de la Loire (CCIPL) for computational resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.643728/full#supplementary-material>

REFERENCES

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: high performance molecular simulations through

multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25. doi: 10.1016/j.softx.2015.06.001

- Arrazola Sastre, A., Luque Montoro, M., Gálvez-Martín, P., Lacerda, H. M., Lucia, A., Llaverro, F., et al. (2020). Small GTPases of the Ras and Rho families

- switch on/off signaling pathways in neurodegenerative diseases. *Int. J. Mol. Sci.* 21:6312. doi: 10.3390/ijms21176312
- Balaji, S., Sujatha, S., Kumar, S. S. C., and Srinivasan, N. (2001). PALI—a database of Phylogeny and ALIGNment of homologous protein structures. *Nucleic Acids Res.* 29, 61–65. doi: 10.1093/nar/29.1.61
- Basse, M. J., Betzi, S., Morelli, X., and Roche, P. (2016). 2P2idb v2: update of a structural database dedicated to orthosteric modulation of protein–protein interactions. *Database* 2016:baw007. doi: 10.1093/database/baw007
- Bendall, C. J., Liu, S., Aumentado-Armstrong, T., Istrate, B., Cernek, P. T., Khan, S., et al. (2014). Transient protein–protein interface prediction: datasets, features, algorithms, and the RAD-T predictor. *BMC Bioinformatics* 15:82. doi: 10.1186/1471-2105-15-82
- Berman, H., Henrick, K., and Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* 10, 980–980. doi: 10.1038/nsb1203-980
- Bielnicki, J. A., Shkumatov, A. V., Derewenda, U., Somlyo, A. V., Svergun, D. I., and Derewenda, Z. S. (2011). Insights into the molecular activation mechanism of the RhoA-specific guanine nucleotide exchange factor, PDZRhoGEF. *J. Biol. Chem.* 286, 35163–35175. doi: 10.1074/jbc.M111.270918
- Chen, Z., Guo, L., Sprang, S. R., and Sternweis, P. C. (2011). Modulation of a GEF switch: autoinhibition of the intrinsic guanine nucleotide exchange activity of p115-RhoGEF. *Prot. Sci.* 20, 107–117. doi: 10.1002/pro.542
- Cheng, T. M. K., Blundell, T. L., and Fernandez-Recio, J. (2007). pyDock: electrostatics and desolvation for effective scoring of rigid-body protein–protein docking. *Proteins* 68, 503–515. doi: 10.1002/prot.21419
- Cherfils, J., and Zeghouf, M. (2011). Chronicles of the GTPase switch. *Nat. Chem. Biol.* 7, 493–495. doi: 10.1038/nchembio.608
- Cherfils, J., and Zeghouf, M. (2013). Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol. Rev.* 93, 269–309. doi: 10.1152/physrev.00003.2012
- Chéron, J. B., Zacharias, M., Antonczak, S., and Fiorucci, S. (2017). Update of the ATTRACT force field for the prediction of protein–protein binding affinity. *J. Comput. Chem.* 38, 1887–1890. doi: 10.1002/jcc.24836
- Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163
- da Silva, A. W. S., and Vranken, W. F. (2012). Acypype-antechamber python parser interface. *BMC Res. Notes* 5:367. doi: 10.1186/1756-0500-5-367
- Derewenda, U., Oleksy, A., Stevenson, A. S., Korczynska, J., Dauter, Z., Somlyo, A. P., et al. (2004). The crystal structure of RhoA in complex with the DH/PH fragment of PDZRhoGEF, an activator of the Ca²⁺ sensitization pathway in smooth muscle. *Structure* 12, 1955–1965. doi: 10.1016/j.str.2004.09.003
- Dvorsky, R., and Ahmadian, M. R. (2004). Always look on the bright site of Rho: structural implications for a conserved intermolecular interface. *EMBO Rep.* 5, 1130–1136. doi: 10.1038/sj.embor.7400293
- Feig, M., and Mirjalili, V. (2016). Protein structure refinement via molecular-dynamics simulations: what works and what does not? *Proteins* 84, 282–292. doi: 10.1002/prot.24871
- Felline, A., Belmonte, L., Raimondi, F., Bellucci, L., and Fanelli, F. (2019). Interconnecting flexibility, structural communication, and function in RhoGEF oncoproteins. *J. Chem. Inform. Model.* 59, 4300–4313. doi: 10.1021/acs.jcim.9b00271
- Gabb, H. A., Jackson, R. M., and Sternberg, M. J. (1997). Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* 272, 106–120. doi: 10.1006/jmbi.1997.1203
- Gasmi-Seabrook, G. M. C., Marshall, C. B., Cheung, M., Kim, B., Wang, F., Jang, Y. J., et al. (2010). Real-time NMR study of guanine nucleotide exchange and activation of RhoA by PDZ-RhoGEF. *J. Biol. Chem.* 285, 5137–5145. doi: 10.1074/jbc.M109.064691
- Guilluy, C. (2010). The Rho exchange factor Arhgef1 mediates the effects of angiotensin II on vascular tone and blood pressure. *Nat. Med.* 16:9. doi: 10.1038/nm.2079
- Hoffman, G. R., and Cerione, R. A. (2002). Signaling to the Rho GTPases: networking with the DH domain. *FEBS Lett.* 513, 85–91. doi: 10.1016/S0014-5793(01)03310-5
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD-visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. doi: 10.1016/0263-7855(96)00018-5
- Ihara, K., Muraguchi, S., Kato, M., Shimizu, T., Shirakawa, M., Kuroda, S., et al. (1998). Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J. Biol. Chem.* 273, 9656–9666. doi: 10.1074/jbc.273.16.9656
- Jiang, J., Wang, N., Chen, P., Zheng, C., and Wang, B. (2017). Prediction of protein hotspots from whole protein sequences by a random projection ensemble system. *Int. J. Mol. Sci.* 18:1543. doi: 10.3390/ijms18071543
- Jiménez-García, B., Pons, C., and Fernández-Recio, J. (2013). pyDockWEB: a web server for rigid-body protein–protein docking using electrostatics and desolvation scoring. *Bioinformatics* 29, 1698–1699. doi: 10.1093/bioinformatics/btt262
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869
- Kortemme, T., Kim, D. E., and Baker, D. (2004). Computational alanine scanning of protein–protein interfaces. *Sci.* 2004:pl2. doi: 10.1126/stke.2192004pl2
- Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., et al. (2017). The ClusPro web server for protein–protein docking. *Nat. Protoc.* 12, 255–278. doi: 10.1038/nprot.2016.169
- Krissinel, E. (2010). Crystal contacts as nature’s docking solutions. *J. Comput. Chem.* 31, 133–143. doi: 10.1002/jcc.21303
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797. doi: 10.1016/j.jmb.2007.05.022
- Kristelly, R., Gao, G., and Tesmer, J. J. G. (2004). Structural determinants of RhoA binding and nucleotide exchange in leukemia-associated Rho guanine-nucleotide exchange factor. *J. Biol. Chem.* 279, 47352–47362. doi: 10.1074/jbc.M406056200
- Lensink, M. F., Brysbaert, G., Nadzirin, N., Velankar, S., Chaleil, R. A. G., Gerguri, T., et al. (2019). Blind prediction of homo- and hetero-protein complexes: the CASP13-CAPRI experiment. *Proteins* 87, 1200–1221. doi: 10.1002/prot.25838
- Lensink, M. F., Méndez, R., and Wodak, S. J. (2007). Docking and scoring protein complexes: CAPRI 3rd edition. *Proteins* 69, 704–718. doi: 10.1002/prot.21804
- Lensink, M. F., Velankar, S., Baek, M., Heo, L., Seok, C., and Wodak, S. J. (2018). The challenge of modeling protein assemblies: the CASP12-CAPRI experiment. *Proteins* 86, 257–273. doi: 10.1002/prot.25419
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins* 78, 1950–1958. doi: 10.1002/prot.22711
- Loirand, G. (2015). Rho kinases in health and disease: from basic science to translational research. *Pharmacol. Rev.* 67, 1074–1095. doi: 10.1124/pr.115.010595
- Loirand, G., and Pacaud, P. (2010). The role of Rho protein signaling in hypertension. *Nat. Rev. Cardiol.* 7, 637–647. doi: 10.1038/nrcardio.2010.136
- Luigia, C. M., Jérémy, B., Nabila, D., Gilliane, C., Anne, B., Michel, A., et al. (2015). Angiotensin II activates the RhoA exchange factor Arhgef1 in humans. *Hypertension* 65, 1273–1278. doi: 10.1161/HYPERTENSIONAHA.114.05065
- Lutz, S., Shankaranarayanan, A., Coco, C., Ridilla, M., Nance, M. R., Vettel, C., et al. (2007). Structure of goq-p63RhoGEF-RhoA complex reveals a pathway for the activation of RhoA by GPCRs. *Science* 318, 1923–1927. doi: 10.1126/science.1147554
- Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The embl-ebi search and sequence analysis tools apis in 2019. *Nucleic Acids Res.* 47, W636–W641. doi: 10.1093/nar/gkz268
- Margreitter, C., and Oostenbrink, C. (2017). MDplot: visualise molecular dynamics. *R J.* 9, 164–186. doi: 10.32614/RJ-2017-007
- Mirjalili, V., Noyes, K., and Feig, M. (2014). Physics-based protein structure refinement through multiple molecular dynamics trajectories and structure averaging. *Proteins* 82, 196–207. doi: 10.1002/prot.24336
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G., Sonnhammer, E. L. L., et al. (2021). Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913
- Mosca, R., Céol, A., and Aloy, P. (2013). Interactome3D: adding structural details to protein networks. *Nat. Methods* 10, 47–53. doi: 10.1038/nmeth.2289
- Ottmann, C. (2015). “New compound classes: protein–protein interactions,” in *New Approaches to Drug Discovery*, Vol. 232, eds U. Nielsch, U. Fuhrmann, and S. Jaroach (Cham: Springer International Publishing), 125–138. doi: 10.1007/164_2015_30

- Petit, A. P., Garcia-Petit, C., Bueren-Calabuig, J. A., Vuillard, L. M., Ferry, G., and Boutin, J. A. (2018). A structural study of the complex between neuroepithelial cell transforming gene 1 (Net1) and RhoA reveals a potential anticancer drug hot spot. *J. Biol. Chem.* 293, 9064–9077. doi: 10.1074/jbc.RA117.001123
- Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., and Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 30, 1771–1773. doi: 10.1093/bioinformatics/btu097
- Prieto-Dominguez, N., Parnell, C., and Teng, Y. (2019). Drugging the small GTPase pathways in cancer treatment: promises and challenges. *Cells* 8:255. doi: 10.3390/cells8030255
- Ramachandran, S., Kota, P., Ding, F., and Dokholyan, N. V. (2011). Automated minimization of steric clashes in protein structures. *Proteins* 79, 261–270. doi: 10.1002/prot.22879
- Rolland, T., Tasan, M., Charlotiaux, B., Pevzner, S., Zhong, Q., Sahni, N., et al. (2014). A proteome-scale map of the human interactome network. *Cell* 159, 1212–1226. doi: 10.1016/j.cell.2014.10.050
- Scheurer, M., Rodenkirch, P., Siggel, M., Bernardi, R. C., Schulten, K., Tajkhorshid, E., et al. (2018). PyContact: rapid, customizable, and visual analysis of noncovalent interactions in MD simulations. *Biophys. J.* 114, 577–583. doi: 10.1016/j.bpj.2017.12.003
- Skwarczynska, M., and Ottmann, C. (2015). Protein–protein interactions as drug targets. *Future Med. Chem.* 7, 2195–2219. doi: 10.4155/fmc.15.138
- Smith, G. R., and Sternberg, M. J. (2002). Prediction of protein–protein interactions by docking methods. *Curr. Opin. Struct. Biol.* 12, 28–35. doi: 10.1016/S0959-440X(02)00285-3
- Sukhwal, A., and Sowdhamini, R. (2013). Oligomerisation status and evolutionary conservation of interfaces of protein structural domain superfamilies. *Mol. Biosyst.* 9, 1652–1661. doi: 10.1039/c3mb25484d
- Takemura, K., and Kitao, A. (2019). More efficient screening of protein-protein complex model structures for reducing the number of candidates. *Biophys. Physicobiol.* 16, 295–303. doi: 10.2142/biophysico.16.0_295
- van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastiris, P. L., Karaca, E., et al. (2016). The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428, 720–725. doi: 10.1016/j.jmb.2015.09.014
- Vetter, I. R. (2014). “The structure of the G domain of the Ras superfamily,” in *Ras Superfamily Small G Proteins: Biology and Mechanisms*, Vol. 1, ed A. Wittinghofer (Vienna: Springer Vienna), 25–50. doi: 10.1007/978-3-7091-1806-1_2
- Wang, W., Yang, Y., Yin, J., and Gong, X. (2017). Different protein-protein interface patterns predicted by different machine learning methods. *Sci. Rep.* 7:16023. doi: 10.1038/s41598-017-16397-z
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., et al. (2018). MolProbity: more and better reference data for improved all-atom structure validation. *Prot. Sci.* 27, 293–315. doi: 10.1002/pro.3330

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Gheyouché, Bagueneau, Loirand, Offmann and Téletchéa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.