

RESEARCH ARTICLE

# Bending-Twisting Motions and Main Interactions in Nucleoplasmin Nuclear Import

Marcos Tadeu Geraldo<sup>1\*</sup>, Agnes Alessandra Sekijima Takeda<sup>1,2</sup>, Antônio Sérgio Kimus Braz<sup>3</sup>, Ney Lemke<sup>1</sup>

**1** Laboratório de Bioinformática e Biofísica Computacional, Departamento de Física e Biofísica, Instituto de Biociências de Botucatu, UNESP – Universidade Estadual Paulista, Botucatu, SP, 18618-970, Brazil, **2** Instituto de Biotecnologia (IBTEC), UNESP – Universidade Estadual Paulista, Botucatu, SP, 18607-440, Brazil, **3** Laboratório de Biologia Computacional e Bioinformática, Centro de Ciências Naturais e Humanas, UFABC – Universidade Federal do ABC, Santo André, SP, 09210-170, Brazil

\* [mtgeraldo@gmail.com](mailto:mtgeraldo@gmail.com)



CrossMark  
click for updates

**OPEN ACCESS**

**Citation:** Geraldo MT, Takeda AAS, Braz ASK, Lemke N (2016) Bending-Twisting Motions and Main Interactions in Nucleoplasmin Nuclear Import. PLoS ONE 11(6): e0157162. doi:10.1371/journal.pone.0157162

**Editor:** Bostjan Kobe, University of Queensland, AUSTRALIA

**Received:** April 7, 2016

**Accepted:** May 25, 2016

**Published:** June 3, 2016

**Copyright:** © 2016 Geraldo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant numbers 2012/19447-2; 2014/21976-9) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (grant number 142110/2012-4). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Alpha solenoid proteins play a key role in regulating the classical nuclear import pathway, recognizing a target protein and transporting it into the nucleus. Importin- $\alpha$  (Imp $\alpha$ ) is the solenoid responsible for cargo protein recognition, and it has been extensively studied by X-ray crystallography to understand the binding specificity. To comprehend the main motions of Imp $\alpha$  and to extend the information about the critical interactions during carrier-cargo recognition, we surveyed different conformational states based on molecular dynamics (MD) and normal mode (NM) analyses. Our model of study was a crystallographic structure of Imp $\alpha$  complexed with the classical nuclear localization sequence (cNLS) from nucleoplasmin (Npl), which was submitted to multiple 100 ns of MD simulations. Representative conformations were selected for calculating the 87 lowest frequencies NMs of vibration, and a displacement approach was applied along each NM. Based on geometric criteria, using the radius of curvature and inter-repeat angles as the reference metrics, the main motions of Imp $\alpha$  were described. Moreover, we determined the salt bridges, hydrogen bonds and hydrophobic interactions in the Imp $\alpha$ -NplNLS interface. Our results show the bending and twisting motions participating in the recognition of nuclear proteins, allowing the accommodation and adjustment of a classical bipartite NLS sequence. The essential contacts for the nuclear import were also described and were mostly in agreement with previous studies, suggesting that the residues in the cNLS linker region establish important contacts with Imp $\alpha$  adjusting the cNLS backbone. The MD simulations combined with NM analysis can be applied to the Imp $\alpha$ -NLS system to help understand interactions between Imp $\alpha$  and cNLSs and the analysis of non-classic NLSs.

## Introduction

Solenoid proteins are molecules composed of structural motifs that are arranged in tandem, creating a superhelical architecture. This modular characteristic provides the establishment of folding and binding contacts that contrasts the globular proteins, allowing higher flexibility and the arrangement of cooperative protein-protein interactions due to the formation of diversified interfaces [1–4].

A remarkable characteristic in nuclear protein import regulation is the contribution of solenoid proteins to recognizing a target protein and transporting it into the nucleus [5]. One of the most studied pathways is the classical nuclear import pathway, which requires the interaction between the solenoid Importin- $\alpha$  (Imp $\alpha$ ) and Importin- $\beta$  (Imp $\beta$ ) proteins, followed by the assembly of the cargo protein to Imp $\alpha$  [6–10]. This trimeric complex is translocated through the nuclear pore complex (NPC), and the cargo protein is delivered into the cell nucleus mediated by Ran-GTP-dependent steps of protein-protein interactions [11, 12]. The dissociation of the cargo protein from Imp $\alpha$  is catalyzed by nucleoporins (e.g., NUP50) [13], and Imp $\alpha$  binds to its export factor CAS complexed with RanGTP [14]. A final step is the recycling of both Imp $\alpha$  and Imp $\beta$  back to the cytoplasm.

Imp $\alpha$  is composed of ten tandem repeats of armadillo motifs (ARM) oriented in an elongated and curved-twisted shape [6, 10, 15]. Each motif is formed by a superhelical architecture of three  $\alpha$ -helices [6, 15, 16]. From the curved orientation, convex and concave surfaces can be identified; in particular, the inner concave surface harbors conserved residues that mediate the cargo protein binding [6, 17].

Cargo protein transport depends on the recognition of a specific sequence signal by Imp $\alpha$  called the nuclear localization sequence (NLS). For classical NLSs (cNLSs), the binding pattern to Imp $\alpha$  is primarily mediated by one or two clusters of positively charged residues, the monopartite or bipartite cNLSs, respectively [18–22]. The clusters of bipartite cNLSs are separated by 10–12 variant residues, denominated as the linker region.

The inner concave surface of Imp $\alpha$  is adapted to receive either monopartite or bipartite cNLSs, and its specific binding sites can be identified as the major and minor sites. Structural and biophysical studies have related important positions of Imp $\alpha$  to the NLS binding: positions P2–P5 from the major site (ARMs 2–4; binding to both monopartite and bipartite cNLSs) and P1'–P2' from the minor site (ARMs 6–8; binding preferentially to bipartite cNLSs) [21, 23]. Moreover, the region between major and minor sites (ARMs 4–6) provides interactions and is also considered fundamental in cNLS recognition [21, 24–30].

Few studies have reported the flexibility and structural integrity of Imp $\beta$  [2, 31]; however, the application of computational simulation approaches is underexplored for understanding the wide structural motions and the interaction basis of nuclear import mediated by Imp $\alpha$  binding. Therefore, two main questions arose: (i) Are there motions related to NLS recognition? (ii) What is the dynamical behavior of the interactions on the complex interface? Motivated by these questions, we combined molecular dynamics simulations (MD) and normal mode (NM) analysis. As our cNLS model of study, we used the crystallographic structure of nucleoplasmin NLS (NplNLS) complexed with Imp $\alpha$  because experimental works involving Npl [17, 24, 32] are available and could be used to support and contrast with our simulation results. Npl is the first protein to be described as a molecular chaperone involved in chromatin reprogramming [33], and it is characterized as containing a bipartite cNLS (Npl:<sup>151</sup>GSAVKRP AATKKAGQAKKKKLD<sup>172</sup>; residues in bold indicate the positions in contact with Imp $\alpha$  minor and major binding sites, respectively). We determined that bending and twisting-like major movements of Imp $\alpha$  may influence the NLS binding. In addition, we confirmed the

importance of contacts in the major and minor sites, along with contacts flanking these sites, including the linker region, for the establishment of carrier-cargo recognition.

## Materials and Methods

### Model of study: Classical bipartite NLS

All simulation analyses were conducted using the crystallographic structure of NplNLS complexed to mouse Imp $\alpha$  isoform 2 (Imp $\alpha$ -NplNLS; PDB ID: 3UL1) [17]. The missing atoms from residues G<sup>152</sup> and S<sup>342</sup> of NplNLS (<sup>152</sup>GSAVKRPAATKKAGQAKKKLD<sup>172</sup>) and Imp $\alpha$  (Imp $\alpha$  $\Delta$ IBB, residues 72–497), respectively, were modeled with MODELLER v9.11 software [34]. The N-terminal Imp $\beta$ -binding (IBB, residues 1 to 71) domain from Imp $\alpha$  was kept truncated because it competes in the binding in the NLS region. The choice criteria for determining the best model were guided by the correct stereochemistry and the occurring interactions in the complex interface, using Molprobit [35] and PISA [36] servers, respectively. An additional modeling for Imp $\alpha$  was conducted removing the peptide (Apo Imp $\alpha$ ) to compare the motions and flexibility of Imp $\alpha$  in the presence and absence of the cNLS.

### MD simulations

The topology and parameter files were generated on the program GROMACS v4.5.3 [37], employing the force field Charmm36 [38] without a protonation requirement because the PROKPA webserver (<http://propka.org>) analysis indicated no protonation changes. A cubic box of 56,568 explicit TIP3P water molecules [39, 40] was generated, ensuring at least 10 Å from the protein system to each edge of the box. Counter ions were added for system-charge neutralization by replacing water molecules. The system was submitted to a gradual energy minimization composed of four steps: (i) 500 steps of energy minimization by the steepest descent method, limiting the protein and peptide movement to accommodate the solvent molecules; (ii) 50,000 steps of energy minimization by the steepest descent method, limiting the movement of the protein and peptide's main chains; (iii) 50,000 steps of unrestricted energy minimization by the steepest descent method; and (iv) 50,000 steps of unrestricted energy minimization by the conjugated gradients (CG) method.

Equilibration and unrestrained MDs were performed in periodic boundary conditions. The leapfrog integrator was used for integrating Newton's equations of motion. The linear constraint solver (LINCS) method [37, 41] was used to freeze bonds involving hydrogen atoms, allowing an integration step of 2 femtoseconds (fs). The cutoff distance for short-range electrostatic and van der Waals interactions was 10 Å. The Particle Mesh Ewald method (PME) [42] was used to treat long-range electrostatics.

The system was equilibrated in two steps, both applying a position restraining force on the heavy atoms of the protein. The first stage involved the adoption of NVT conditions (constant number of particles, volume and temperature), heating the system to the target temperature of 300 kelvin (K) and simulating in this condition by 100 picoseconds (ps) with a velocity-rescaling (V-rescale) thermostat [43]. The second stage involved 100 ps of equilibration by the adoption of NPT conditions (constant number of particles, pressure and temperature), with pressure coupling using the Parrinello-Rahman barostat [44] and keeping the pressure relatively constant, close to the value of 1 bar. After the equilibration procedures, the restraints were removed, and the system was submitted to three MD simulations of 100 nanoseconds (ns) each, with structure sampling every 10 ps.

To ensure the conformational sampling of the system, a clustering technique proposed by Lyman and Zuckerman [45] was applied. In this approach, the steps below were applied to generate a collection of reference structures of the simulations  $\{S_i\}$ : (1) A cutoff root-mean square

deviation (RMSD)  $d$  was defined. (2) Each simulation was merged into a single trajectory file. (3) One structure from the trajectory file was sampled randomly and denominated as the reference structure. (4) All structures were compared to the sampled reference structure, and the ones with an RMSD less than  $d$  were removed from the trajectory file. (5) Steps 1, 2, 3 and 4 were repeated until all structures were removed, thus generating a collection of reference structures  $\{S_i\}$ . (6) Based on the collection  $\{S_i\}$ , each frame from the trajectory file was clustered with the nearest reference structure, and the frequency of structures of each cluster was then calculated. An estimation of convergence is assessed when each reference structure is equivalently represented in each simulation. The parameter for the calculation of  $d$  was the global RMSD of carbon- $\alpha$  ( $C\alpha$ ) atoms. The choice of  $d$  ( $= 2.5 \text{ \AA}$ ) was constrained to a feasible number of reference structures for the subsequent analysis of NMs.

## NM analysis

Each reference structure obtained from MD simulations was minimized in the GROMACS program with explicit solvent using the same methodology described above. All NM analyses were carried out in the CHARMM v.36b1 program [46]. The topology and parameter files required for CHARMM were generated with the CHARMM-GUI server ([www.charmm-gui.org](http://www.charmm-gui.org)) employing an additional energy minimization. The CG algorithm was applied with harmonic constraints that were progressively decreased from 250 to 5 kcal/mol $^{-1}\text{\AA}^{-2}$ , with 100 steps of minimization at each decrease. Then, the constraints were removed, and 10,000 steps of CG were carried out. Afterwards, the adopted basis Newton Raphson (ABNR) algorithm was applied with no constraints for 300,000 steps.

The final minimized structure was used for the calculation of the 87 lowest frequency NMs, using the VIBRAN module of CHARMM for each reference structure. An NM-displacement method using the VMOD facility of CHARMM was applied, generating structures along each NM based on short MD simulations at a low temperature (30 K), followed by energy minimization. Based on the values of mass-weighted root mean square (MRMS), the maximum displacement range was set to 3  $\text{\AA}$  for each direction of the NM with a 0.1  $\text{\AA}$  projection step, totaling 61 structures per NM. For each MRMS step, a harmonic force constant over the  $C\alpha$  atoms was applied (increasing from 1,000 until 10,000 kcal/mol $^{-1}\text{\AA}^{-2}$ ), and a short MD simulation was carried out for 1 ps for each constant value, totaling 10 ps of simulation. Keeping the restraints, 1,000 steps of CG energy minimization were employed to generate the final structure. In addition, for each projection step, a value of the total restraint energy, according to the miscellaneous mean field potential (MMFP) facility of CHARMM, was used as a criterion for discarding unfavorable conformations.

## Data Analysis

**PCA over MD simulations.** Principal component analysis (PCA) from MD simulations was performed using the quasi-routine of the module VIBRAN of CHARMM to obtain the covariance matrix of  $C\alpha$  atomic displacements from the trajectories and identify the most relevant structural variations. The calculation was performed according to the description of Floquet et al. [47], and the resulting principal components (PCs) were used to compare the motions observed in NMs.

**Collectivity.** The measurement of the involvement of atoms in a particular protein motion (referred to as degree of collectivity) for a given NM was calculated according to Bruschiweiler [48] and Tama and Sanejouand [49], using an *in-house* CHARMM script. The degree of collectivity is comprised between 0 and 1. Values close to 1 indicate maximum collectivity.

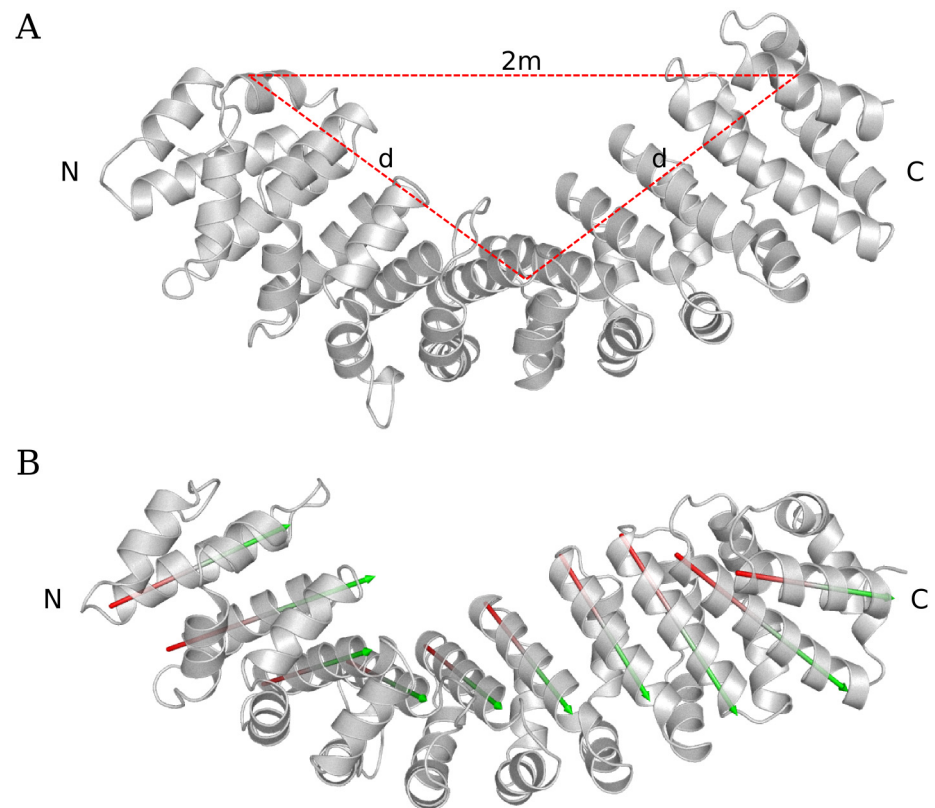
**Geometric analysis of Imp $\alpha$ .** The wide motions from Imp $\alpha$  were described in terms of geometrical measurements. For bending characterized motions, three vertices (represented by Imp $\alpha$  residues R<sup>117</sup>, A<sup>313</sup> and K<sup>486</sup>) were manually selected: two vertices in distal tips and one vertex at the middle point of the protein. The plane formed by these three vertices corresponded to the plane of the observed bending motion (Fig 1). Then, the radius of curvature ( $R$ ) was calculated:

$$R = \frac{d^2}{2\sqrt{d^2 - m^2}} \quad (1)$$

$d$  is the distance from the distal vertices to the middle vertex, and  $m$  is half the distance between the two distal vertices.

For twisting characterized motions, the angle between  $\alpha$ -helices 3 (H3) from neighboring ARM repeats were calculated using an available script for PyMol (<http://www.pymolwiki.org/index.php/AngleBetweenHelices>). For each pair of helices, vectors were defined along the C $\alpha$  atoms, and the torsion angles between these vectors were determined.

This geometric analysis was also applied to X-ray-solved structures of Imp $\alpha$  complexed with different types of bipartite NLSs. These structures were retrieved by the basic local alignment search tool allocated at Prody software [50] using the Imp $\alpha$  -NplNLS model as the query. Only structures with 100% of sequence identity to Imp $\alpha$  were selected to be compared to the simulation data.



**Fig 1. Scheme of geometrical methods adopted for the description of the Imp $\alpha$  motions.** (A) Display of the selected vertices and the determined distances (2m and d) for the calculation of the radius of curvature. (B) The vectors generated to H3 for each ARM repeat.

doi:10.1371/journal.pone.0157162.g001



**Maps of cross-correlations.** The initial comparison of residue-residue contacts from the MD and NM results was initially performed by the generation of maps of cross-correlation to evaluate the associated movements between NplNLS and Imp $\alpha$ . The calculation of correlations corresponded to the ensemble of the trajectories of all MD simulations into one final pseudo-trajectory. Similarly, an ensemble of all structures from NM-displacements into one pseudo-trajectory was conducted. The cross-correlation calculations were performed in the Wordom software [51].

**Interactions evaluation.** The occurrence of specific interactions in the Imp $\alpha$ -NplNLS interface was evaluated. The determination of salt bridges and hydrogen bonds were performed using the VMD software [52]. The criteria for considering the occurrence of these interactions were the donor-acceptor distance for salt bridges and hydrogen bonds  $\leq 3.5$  Å and the donor-hydrogen-acceptor angle deviation for hydrogen bonds  $\leq 60$  degrees. The determination of hydrophobic contacts was based initially on the LIGPLOT program, using the crystallographic and *in silico* model of Imp $\alpha$  -NplNLS to generate a list of possible interactions. Later, based on this list, the distances of the closest carbon atom from the hydrophobic side chains of each residue-pair were calculated for the MD and NM ensemble pseudo-trajectories to calculate the percentage of the occurrence of each hydrophobic contact. The criterion adopted was distances  $\leq 4$  Å.

**Complementary analysis.** The backbone RMSD and the C $\alpha$  root-mean-square fluctuations (RMSF) of MD and NM-displacements were calculated in the Wordom software [51]. The structural alignment of NplNLS peptides was performed in the Crystallographic Object-Oriented Toolkit (Coot) software [53]. All graphics from the calculations above were performed using the R software [54], and the structural analysis, visualization and generation were performed in the Pymol software [55].

## Results

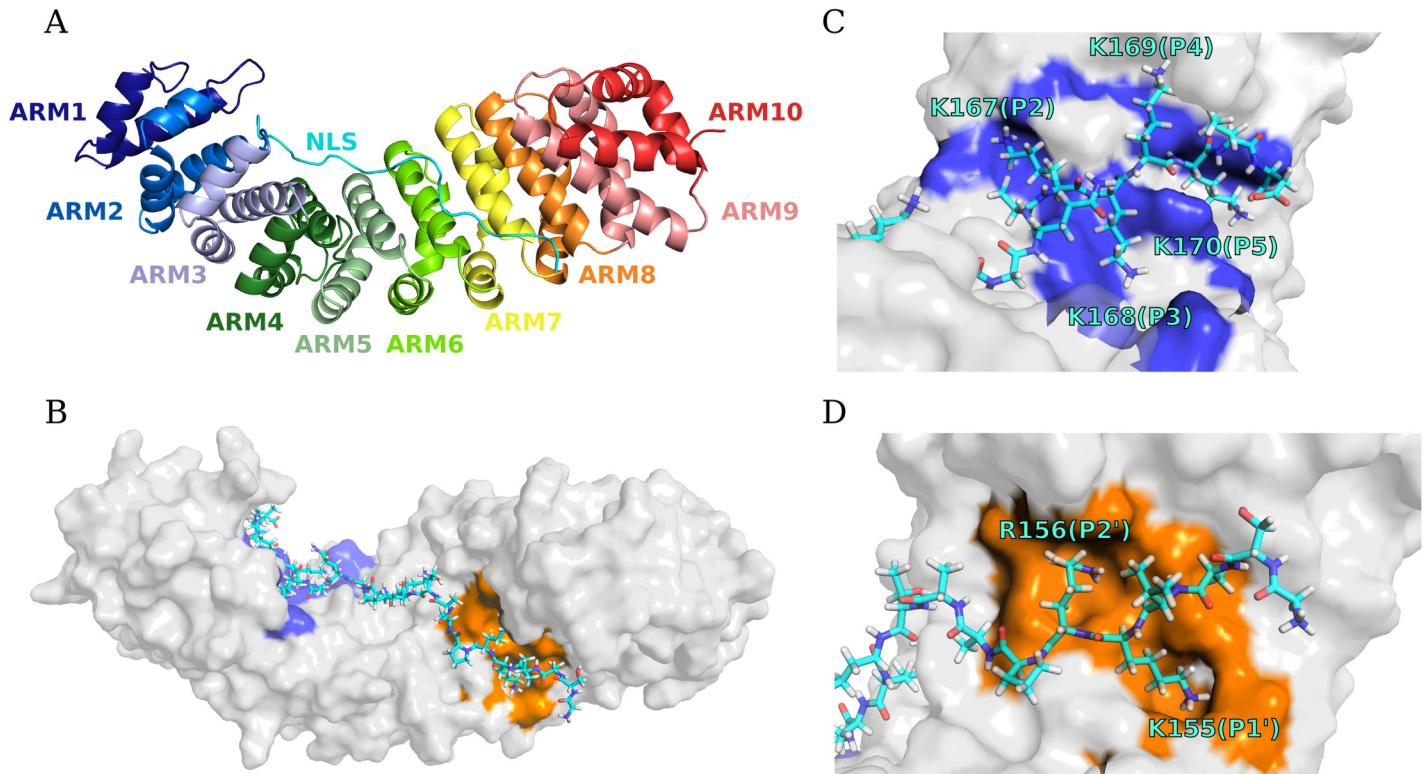
### Selection of Imp $\alpha$ -NplNLS model

A stereochemical analysis allowed for the selection of the best complex model, taking into account the geometry and maintenance of the main interactions in the peptide-protein interface. The best-selected model showed a Ramachandran plot for Imp $\alpha$  with 96.7% and 3.3% of the residues in favored and allowed regions, respectively, without residues in the outlier region, whereas NplNLS had 100% of the residues in favored regions.

An overall evaluation of the Imp $\alpha$  -NplNLS model indicated the NplNLS harboring in both major and minor binding sites of Imp $\alpha$ , with the main residues inside the binding pockets (Fig 2). Furthermore, we were able to evidence the previously stated interactions in the interface of this complex [17, 24]; the only major exception was the absence of Imp $\alpha$ D<sup>325</sup> in contact with NplK<sup>155</sup> in the minor site (S1 Fig).

### Standard MD combined with NM-displacement method

The trajectories obtained from the three Imp $\alpha$ -NplNLS MD simulations were clustered into three reference structures (67,730 ps, 207,080 ps and 274,970 ps), each one exhibiting a similar frame-frequency among simulations, indicating a likely convergence of the MDs (S2 Fig). Reference structure 207,080ps was the most representative in the simulations (approximately 80% of the trajectories clustered with this structure) and was more similar to the X-ray-solved NplNLS structure based on the backbone RMSD values (S1 Table) obtained from the structural alignment of the NLSs (S3 Fig). In this alignment, important positions from major and minor sites were occupied by the expected residues, with similar side chain conformations, represented by NplK<sup>167</sup>, NplK<sup>168</sup> and NplK<sup>170</sup> in positions P2, P3 and P5, respectively, and NplK<sup>155</sup> and



**Fig 2. The starting structure of Imp $\alpha$ -NpINLS for MD simulations.** (A) The Imp $\alpha$  as a cartoon diagram colored based on each ARM repeat as a rainbow spectrum from N-terminal (blue) to C-terminal (red) and the NpINLS as a cyan cartoon diagram positioned in an antiparallel configuration compared to Imp $\alpha$ . (B) The surface representation of Imp $\alpha$  with the NpINLS as a cyan stick diagram, indicating both major (blue) and minor (orange) binding sites. (C) The major site zoom indicating positions P2–P5 and (D) the minor site zoom in P1' and P2'. In both sites, the positively charged side chains are positioned in the main pockets of the Imp $\alpha$  binding core.

doi:10.1371/journal.pone.0157162.g002

NpI<sup>R156</sup> in P1' and P2', respectively. Reference structures 67,730 ps and 274,970 ps showed a greater structural variance of side chains in the major site than in the minor site.

The subsequent step was the NM-displacement method applied to the three MD reference structures. We observed flexible and favorable motions in the early NMs (modes 7–20) of Imp $\alpha$ -NpINLS, represented by lower values of restriction energy along the whole displacement range, compared to the remaining modes (blue areas in the [S4 Fig](#)). Similar patterns were also observed for Apo Imp $\alpha$ ; however, the favorable motions were extended up to NM29 ([S5 Fig](#)), indicating that more conformations could be acquired favorably compared to the Imp $\alpha$  bound to NpINLS.

The distribution of backbone RMSD showed that standard MD in combination with the NM-displacement method increased the conformational exploration, reaching values over 5 Å, particularly to reference structures 207,080 ps and 274,970 ps, whereas MD alone reached only approximately 2.5 Å ([S6 Fig](#)). Moreover, greater C $\alpha$  fluctuations were observed in the NM-displacement technique, based on the RMSF values ([S7](#) and [S8 Figs](#)). The comparison between the RMSF values for Imp $\alpha$ -NpINLS and Apo Imp $\alpha$  showed small differences, primarily limited to the region of the major site (within the range of residue 100 to 200), where Apo Imp $\alpha$  was more flexible ([S9 Fig](#)).

## Collective motions of Imp $\alpha$

The description of the motions obtained from MD and NM calculations combined the results from a qualitative vector-based analysis and a quantitative geometrical analysis. The first NMs showed wide and collective types of motions of the ARM repeats of Imp $\alpha$ , primarily along modes 7–17 (S10 Fig). A qualitative analysis of vectors from modes 7 and 9 clearly showed a motion pattern, described by a bend and a twist (Fig 3 and S1 and S3 Movies), respectively.

The bend motions in NM7/PC1 were characterized by the opening and closing of Imp $\alpha$  in the concave surface, along the NLS binding pockets. The quantitative analysis using geometric measurements determined the radius of curvature in opened and closed configurations of Imp $\alpha$  (Fig 4A, S12 Fig and S1 Movie) and showed different amplitudes for the bending motion, in which Apo Imp $\alpha$  had higher amplitudes than the Imp $\alpha$ -NplNLS complex.

The twist pattern for the Imp $\alpha$  structures along NM9/PC3 (Fig 3, S12 Fig and S3 Movie) oscillated from maximum and minimum values of torsion over the entire protein (Fig 4B). A general observation of the angles for the Imp $\alpha$ -NplNLS complex and Apo Imp $\alpha$  showed similarities among the structures. A more detailed analysis, considering the inter-repeat angles (Fig 5), detected oscillation between the ARMs in the Imp $\alpha$ -NplNLS complex, primarily between ARMs 5–6. Movements with smaller amplitudes were observed for most pairs of ARMs in Apo Imp $\alpha$ , except the oscillation observed between ARMs 6–7.

Bending and twisting movements were also evaluated for the crystal structures of Imp $\alpha$  in the presence of different types of bipartite cNLSs. A comparison of geometries indicated a small difference among them in the order of tenths of ångströms (S5 Table). The qualitative vector analysis of Imp $\alpha$ -NplNLS NM8/PC2 indicated a combination of two motions, characterized by a “lateral” bending tendency (based on a 90° X-axis rotation of Imp $\alpha$  in relation to the bending orientation in NM7/PC1) in ARMs 1–6, mixed with a twist in ARMs 7–10 (Fig 3, S12 Fig and S2 Movie). The amplitude of the inter-repeat angle variation was slightly distinct between NM8 and PC2; however, in both cases, we could verify motions that almost reached both sites, including their intermediate region (Fig 5). The complexity of these mixed motions increased in the following modes (S11 and S13 Figs). The same analysis for Apo Imp $\alpha$  NM8 showed the reduction of movement amplitude for most ARMs.

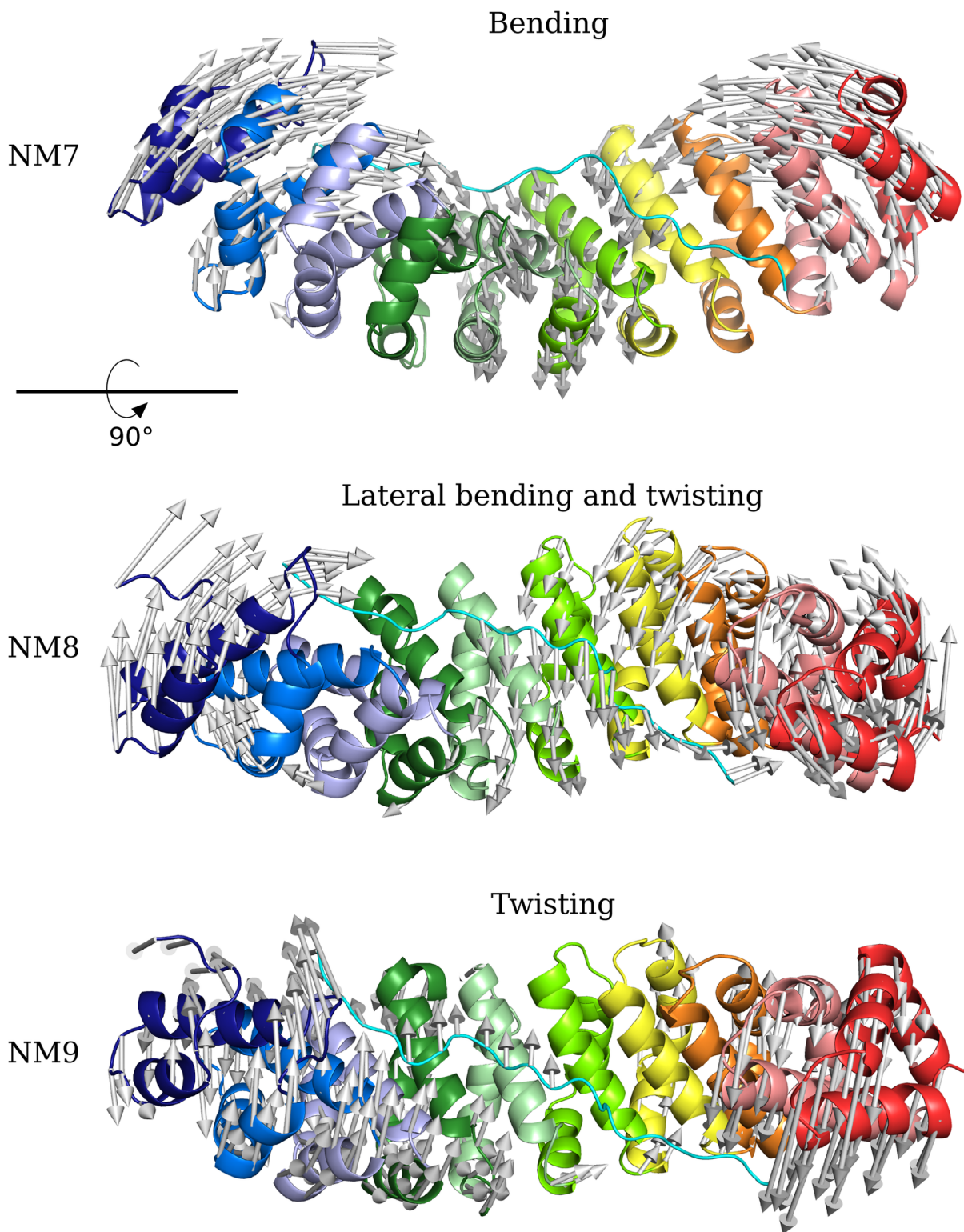
## Main contacts in Imp $\alpha$ -NplNLS interface

In general, the cross-correlations results from MD and NM analyses were similar (Fig 6). Well-bounded areas of positive correlations could be identified, highlighting the NplNLS range of residues in contact with Imp $\alpha$  major and minor sites and the linker region. However, for standard MD, the correlations were more scattered in the linker.

Subsequent analysis showed that most correlations could be related to specific interactions in the interface of Imp $\alpha$ -NplNLS. Compared to our simulation data, the starting structure exhibited more contacts of hydrogen bonds and hydrophobic interactions (S1 Fig). Throughout the simulations, some of those initial contacts were lost, and MD and NM showed mostly common results (Fig 7). In particular, we observed salt bridges (Imp $\alpha$ E<sup>396</sup>, Imp $\alpha$ D<sup>280</sup>, Imp $\alpha$ D<sup>192</sup> and Imp $\alpha$ D<sup>270</sup>) that were established along the NplNLS, specifically in P2', P2 and the linker region. Hydrophobic contacts mediated by tryptophans (Imp $\alpha$ W<sup>399</sup>, Imp $\alpha$ W<sup>357</sup>, Imp $\alpha$ W<sup>273</sup>, Imp $\alpha$ W<sup>231</sup>, Imp $\alpha$ W<sup>184</sup> and Imp $\alpha$ W<sup>142</sup>) occurred mostly in major and minor sites, specifically in P2', P3 and P5. As expected, a great number of hydrogen bonds were established along the NplNLS, such as Imp $\alpha$ S<sup>360</sup>, Imp $\alpha$ N<sup>361</sup>, Imp $\alpha$ G<sup>323</sup>, Imp $\alpha$ V<sup>321</sup>, Imp $\alpha$ R<sup>315</sup>, Imp $\alpha$ Y<sup>277</sup>, Imp $\alpha$ R<sup>238</sup>, Imp $\alpha$ A<sup>148</sup>, Imp $\alpha$ G<sup>150</sup> and Imp $\alpha$ N<sup>188</sup>.

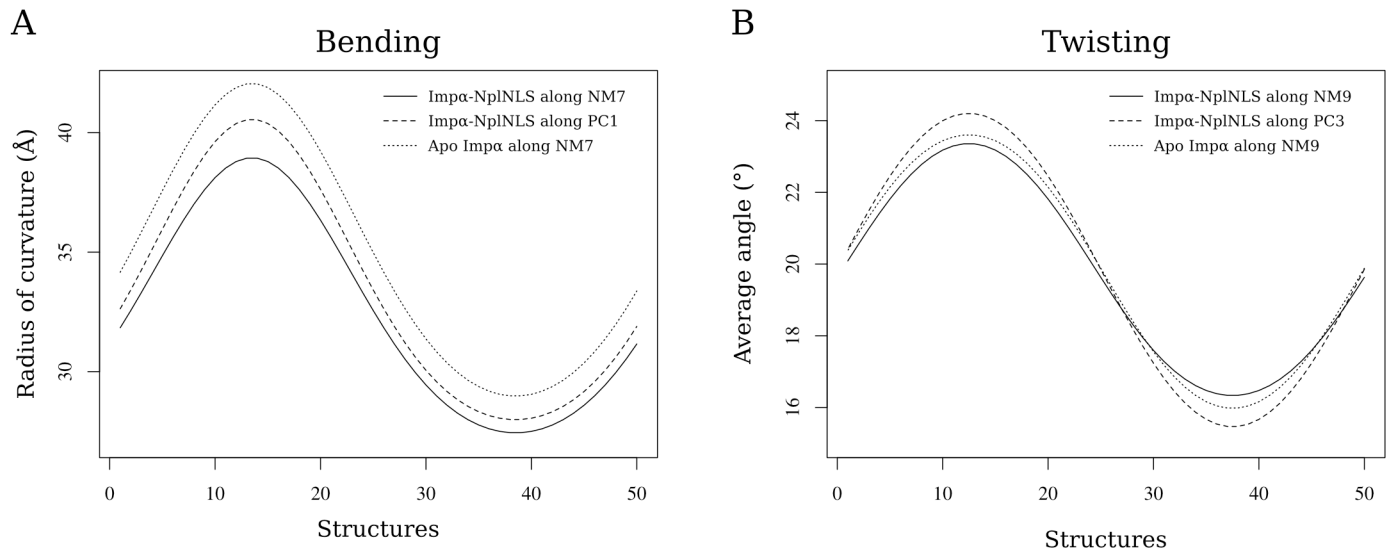
Most of the aforementioned interactions occurred with side chains of the charged residues of the NplNLS. Moreover, we highlighted residues Imp $\alpha$ D<sup>192</sup>, Imp $\alpha$ E<sup>396</sup>, Imp $\alpha$ W<sup>184</sup>, Imp $\alpha$ W<sup>231</sup>,





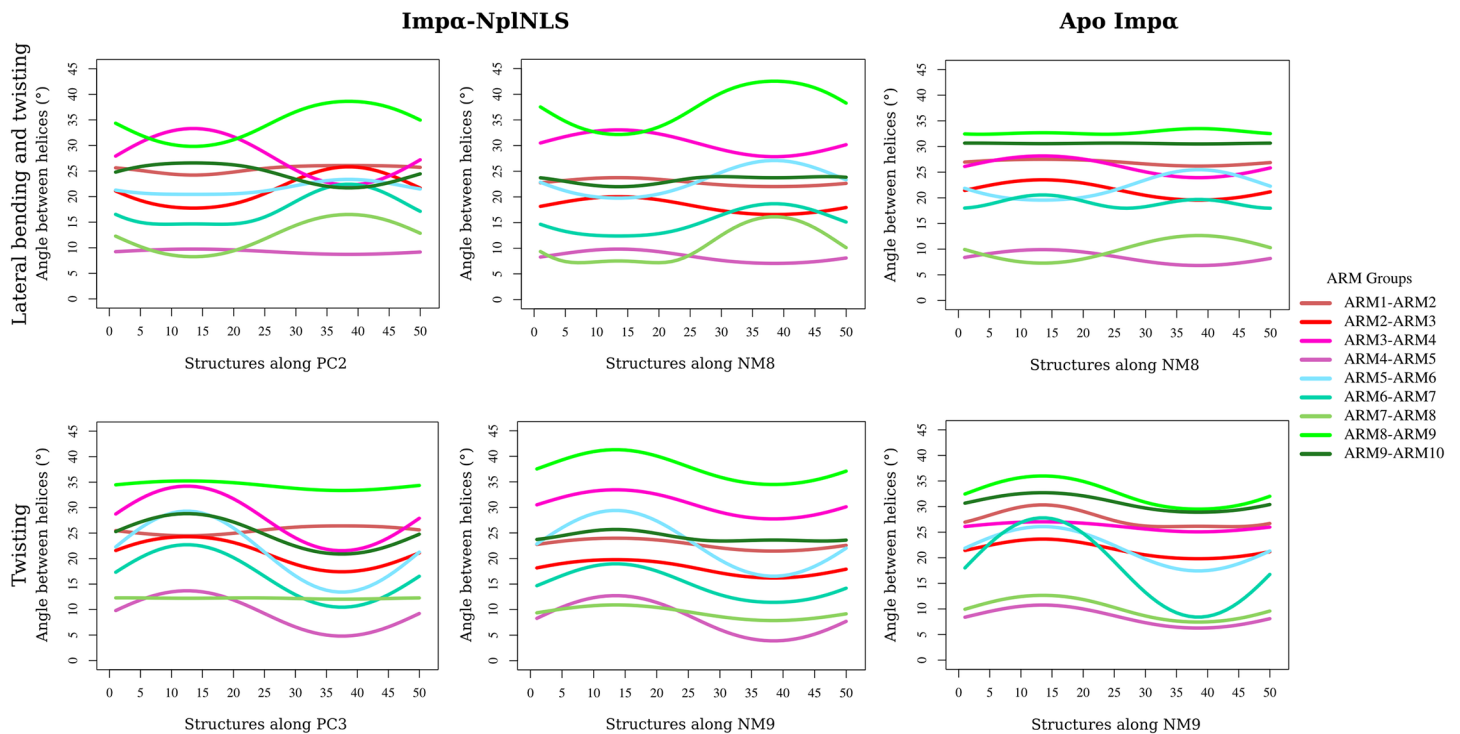
**Fig 3. Main motions observed from NM analysis of Imp $\alpha$  - NpINLS.** Imp $\alpha$  (cartoon model) is shown as a rainbow spectrum from N-terminal (blue) to C-terminal (red), and NpINLS (cyan cartoon model) is positioned in an antiparallel configuration compared to Imp $\alpha$ . The vector arrows for NM7–9 are shown with the correspondent description of the motion that they described. NM7 is shown in a front view, whereas NM8 and NM9 are shown in an upper view (90° rotation in the X-axis).

doi:10.1371/journal.pone.0157162.g003



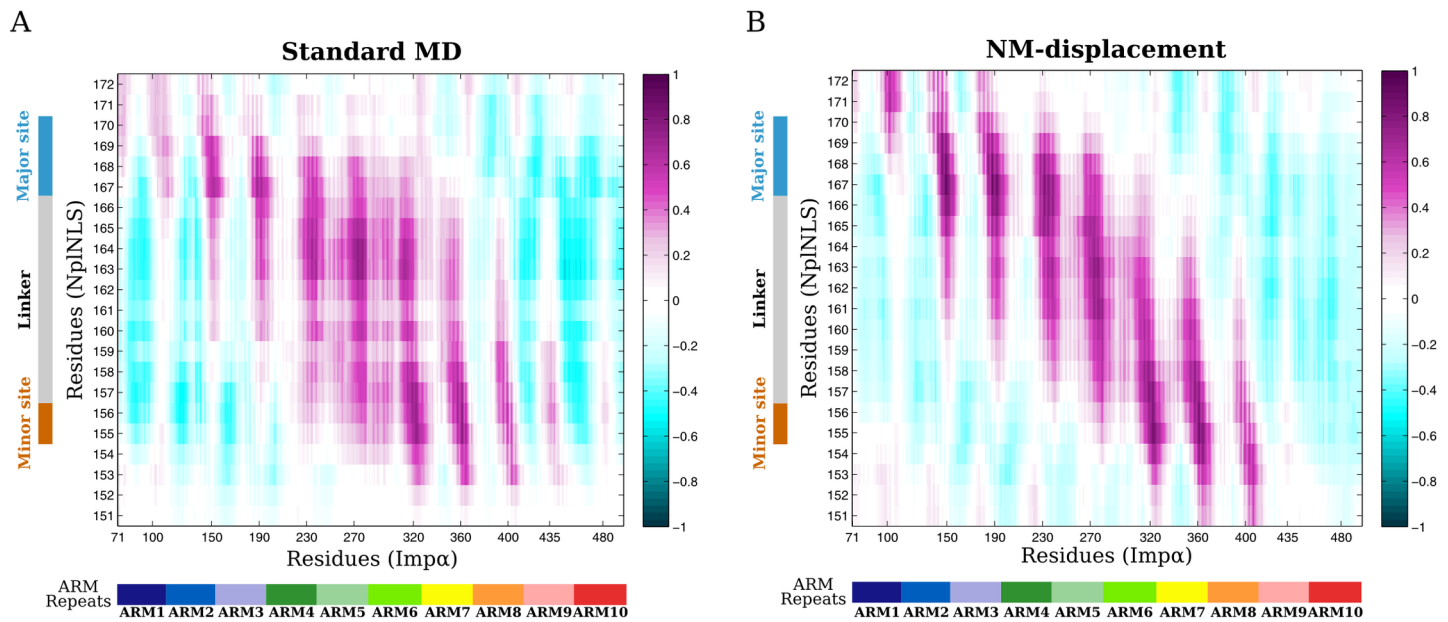
**Fig 4. Geometric analysis of the bending and twisting motions of Imp $\alpha$ .** (A) The bending motion was quantitatively characterized by the radius of curvature along NM7 (solid line) and PC1 (dashed line) for Imp $\alpha$  -NpINLS and along NM7 for Apo Imp $\alpha$  (dotted line), whereas (B) the twisting motion was quantitatively characterized by the average values for the angles between helices along NM9 (solid line) and PC3 (dashed line) for Imp $\alpha$  -NpINLS and along NM9 for Apo Imp $\alpha$  (dotted line).

doi:10.1371/journal.pone.0157162.g004



**Fig 5. Angle between helices of Imp $\alpha$ .** The angles between neighboring H3 pairs from the motions described as lateral-bending/twisting (along PC2 and NM8) and twisting (along PC3 and NM9) for Imp $\alpha$  -NpINLS and Apo Imp $\alpha$ . The ARM groups considered for each angle calculation are depicted with different color assignments.

doi:10.1371/journal.pone.0157162.g005



**Fig 6. Heatmap of cross-correlations between *Impα* and NpINLS.** (A) Trajectories from standard MD simulations (300 ns ensemble) and (B) NM-displacement (ensemble from references structures 67,730 ps, 207,080 ps and 274,970 ps) were used for the calculation of correlations. A color bar indicates the degree of correlation from anti-correlated (negative values) to correlated (positive values) residues. The X and Y axes, respectively, show the position of each ARM repeat in relation to the *Impα* sequence and the NpINLS residues in contact with the protein binding sites.

doi:10.1371/journal.pone.0157162.g006

$Imp\alpha W^{357}$ ,  $Imp\alpha W^{399}$  and  $Imp\alpha Y^{277}$  that were observed in more than 90% of the analyzed trajectory frames from MD and NM ensembles (S2, S3 and S4 Tables).

## Discussion

### Bending and twisting motions may be directly related to *Impα* function

Both bending and twisting motions are good candidates to adapt the *Impα* to the cNLS, because their motion pattern promotes conformational changes in the NLS binding pockets. The combination of these motion patterns appeared to be similarly recurring in other modes, such as NM8. The vectors from other high-collectivity modes, such as NM10–13 and NM17–18, also exhibited these two main motions but in smaller portions of the protein, and they were also associated with undefined types of motions. This finding was expected because modes with higher frequency are normally associated with localized vibrations [56]. Studies involving the protein *Impβ*—a solenoid protein similar to *Impα* [1, 31, 57]—suggested the importance of bending and twisting motions to generate the flexibility of *Impβ*, allowing it to bind to different types of proteins [31, 58]. In our *Impα* computational analysis, the described movements could be equally important to adapt to different sizes of cargo proteins and NLSs, enhancing the contacts over the NLS binding site. Although no significant changes were observed by the geometric analysis from the *Impα* crystallographic structures bound to different types of bipartite NLSs, we must consider that the presence of only a peptide (approximately 20 amino acid length) submitted to similar crystallization conditions may not be sufficient to induce large conformational changes in this system. Therefore, the role of *Impα* motions is likely critical for the cNLS accommodation, considering the entire protein that contains it.

The MD/NM approaches complementarily showed the flexibility of *Impα* by the analysis of NM7–9 and PC1–3. The bending observed in NM7/PC1 may be directly related to the accessibility to the binding sites and the release of NLSs. This movement was observed in both





contacts are depicted in the scheme as black sticks. The main chain of the NplNLS is represented as a gray horizontal line with its respective amino acid sequences, together with side chains shown as perpendicular lines. Only interactions that had an occupancy rate  $\geq 50\%$  of the analyzed trajectories are indicated in the scheme.

doi:10.1371/journal.pone.0157162.g007

Imp $\alpha$ -NplNLS and Apo Imp $\alpha$ , with small differences in their amplitude, resembling “open” and “close” movements. Bending motions have been reported in globular proteins to identify movements of domains along with opened and closed states [59, 60]. The bending analysis for Imp $\alpha$  was possible because its small curvature [1] allowed for the establishment of a plane comprising two distal (N- and C-termini) atoms and a central atom. Solenoid Imp $\beta$  required a more complex analysis, determining angles between vectors projected onto a reference plane in each motif to evaluate the curvature changes [31]. Specifically, the concerted bending motion of Imp $\alpha$  NM7 could operate as an opening-closing gateway, allowing the NLS entrance and adjusting the amplitude of this motion in relation to the cargo protein size. The absence of a ligand may imply in a wider curvature favoring the access to the inner concave surface of Imp $\alpha$ , allowing the IBB domain or cargo protein binding. Pumroy et al. [61] compared the flexibility of three human Imp $\alpha$  isoforms ( $\alpha$ -1,  $\alpha$ -3 and  $\alpha$ -7) considering their bound and unbound states and applying MD simulations. Based on the protein end-to-end distance measurements, the authors observed an increase in the flexibility of Apo Imp $\alpha$  isoforms, in accordance with our bending characterization that showed a higher radius of curvature for Apo Imp $\alpha$ .

According to Kobe et al. [1], the “twist” takes into account the rotations of the neighboring repeats relative to each other along the backbone direction, and ARM repeat proteins have large twist movements, allowing the accommodation of extended and flexible peptides. The twist observed in the NM9/PC3 of Imp $\alpha$ -NplNLS showed similarities to each other, confirming the observations of Hayward et al. [62], which compared NM/PC for lysozyme protein. Although Imp $\alpha$ -NplNLS and Apo Imp $\alpha$  had similar average values (Fig 4B), the differences were found only in individual analyses of the neighboring ARM repeats. The angle variation between ARM6 and ARM7 in NM9 of Apo Imp $\alpha$  (Fig 5) is remarkable, and a similar result was already observed for Apo Imp $\alpha$ -3 [61]. The lack of a ligand (IBB domain or NLS) interacting with the Apo Imp $\alpha$  concave portion may lead to a higher twist in the middle of Apo Imp $\alpha$  due to the lack of interactions that stabilize and provide binding specificity. The increment of the angle amplitude in this region was contrasted by lower values for the remaining protein regions compared to Imp $\alpha$ -NplNLS. Therefore, Imp $\alpha$ -NplNLS may require conformational changes for the NLS adjustment to the binding pockets of the Imp $\alpha$  inner surface, which explains the greater variation in the ARM's angles in contrast to Apo Imp $\alpha$ , which would not require such adjustments. This same explanation applies to the amplitude variations also found in NM8 for the bound/unbound states of Imp $\alpha$ .

Experiments and MD simulation with *Neurospora crassa* Imp $\alpha$  emphasize the instability of an N-terminally truncated Imp $\alpha$  in the absence of an NplNLS peptide, indicating the occupancy of the NLS binding sites as a requirement for Imp $\alpha$  crystallization [63]. In addition, Falces et al. [64] also showed, performing circular dichroism assays with Imp $\alpha$ -1 $\Delta$ IBB from *Xenopus laevis*, the stabilization of the protein upon association with Npl, thus reinforcing the aforementioned observations. The establishment of polar contacts settles the NLS backbone, whereas hydrophobic and electrostatic interactions with positively charged NLS residues allow the specificity [16, 21]. The greater structural flexibility of Apo Imp $\alpha$  was only observed by the geometrical analysis for the NM7 bending. Despite the non-association of the NM8/NM9 twisting motions to a greater overall flexibility of Imp $\alpha$ , the NLS absence allowed for a greater motion amplitude of other NMs, such as modes 21–27, comparing the restriction energies of Imp $\alpha$ -NplNLS (S4 Fig) and Apo Imp $\alpha$  (S5 Fig), indicating that the higher flexibility of Apo Imp $\alpha$  appeared to be



distributed over these modes. In addition, considering the  $C\alpha$  fluctuations of  $\text{Imp}\alpha$  in its bound and unbound states (S9 Fig), we observed a higher flexibility in the major site of Apo  $\text{Imp}\alpha$ . This result was also observed for Apo  $\text{Imp}\alpha$ -3 [61], and the authors predicted a weaker binding for NLSs that relies primarily on the major site. However, this statement does not seem to apply to our case because  $\text{Imp}\alpha$ -2 has been crystallized with different types of NLSs, including those with preference to the major site; e.g., simian virus 40 (SV40) NLS [24]. Moreover, the binding stability of  $\text{Imp}\alpha$ -2 to some of those NLSs was also observed with ligand binding assays [17, 65, 66]. In summary, the NLS binding stabilizes  $\text{Imp}\alpha$ , apparently restricting the motion range represented by some higher-frequency NMs; however, it may also allow localized adjustments between the ARM repeats to improve the overall affinity to the bound NLS.

### The role of the linker residues in cNLS recognition

The cross-correlation calculations performed for both MD simulations (Fig 6A) and displacement along NMs (Fig 6B) showed positive correlations from residues of both major and minor sites of  $\text{Imp}\alpha$ . Further analysis showed that the majority of positive correlations observed could be explained by the correspondence of salt bridges, hydrogen bonds and hydrophobic interactions. We could determine that positions P2 and P5 from the major site and P1' and P2' of the minor binding site have high levels of contact with the NplNLS, based on the occupancy values of some interactions in those regions. These data confirm the maintenance of the main contacts between  $\text{Imp}\alpha$  and NplNLS, showing that our simulation results are in agreement with the experimental data [17, 24].

A classical bipartite NLS sequence interacts simultaneously with major and minor binding sites of  $\text{Imp}\alpha$  and depends on a linker region containing a minimum of 10 residues between P2' and P2 positions to act as a cNLS [5, 30]. The linker region in the NplNLS appears to play a key role in the process of cNLS recognition. The cross-correlation maps clearly indicated an involvement of this region in the promotion of contacts with  $\text{Imp}\alpha$ . A detailed analysis of the interactions showed the occurrence of salt bridges, hydrogen bonds and hydrophobic contacts along the NplNLS. The involvement of  $\text{Imp}\alpha\text{R}^{238}$ ,  $\text{Imp}\alpha\text{R}^{315}$ ,  $\text{Imp}\alpha\text{W}^{273}$  and  $\text{Imp}\alpha\text{Y}^{277}$  was already observed in  $\text{Imp}\alpha$ -NplNLS crystal structures [17, 24]. However, the computational approach indicated other interactions previously reported in single structures;  $\text{Imp}\alpha\text{K}^{353}$  and  $\text{Imp}\alpha\text{N}^{350}$  were reported with the NLS from FEN-1 [30] and  $\text{Imp}\alpha\text{D}^{280}$  from SV40 NLS bound to the  $\text{Imp}\alpha$  from a filamentous fungus [67]. Moreover, new possible contacts were detected; e.g., residues  $\text{Imp}\alpha\text{N}^{446}$ ,  $\text{Imp}\alpha\text{R}^{101}$  and  $\text{Imp}\alpha\text{R}^{102}$  interacted in the N- and C-terminal regions of the NLS.

Our simulation analyses suggest that the linker contacts are important to settle the cNLS, and help to accommodate the cNLS side chains into the grooves of major and minor binding sites. The linker region, in addition to the N- and C-terminals of an NLS, may also compensate for interactions for the establishment of an activation pattern of the other NLS region [22], which indicates that the linker contacts may occur in different NLSs and can be maintained after the docking of the bipartite NLS in both major and minor binding sites, depending on the residues composing the linker, such as proline and acidic amino acids [5, 22]. In summary, the recurrent presence of residues outside the major and minor sites strongly reinforces their importance for the proper binding of the NLS, which corroborates other studies involving other NLSs [21, 25–29] and further encourages us to understand in greater detail their roles and effects during the nuclear import process.

### NM analysis with classic MD simulations

MD simulations and NM analysis have been used for macromolecules as a complementary analysis to the experimental data to describe their main motions and relate to a specific

function [59, 60, 68, 69]. The approach used for the Imp $\alpha$ -cNLS complex combines NMs robustness—to evaluate wider protein movements—with the reference structures from the MD and indicates the benefits of this association in protein-peptide analysis.

The data sampling from MD simulations, despite the time constraints and convergence difficulties associated with this technique [70, 71], were balanced among the selected reference structures. Considering the time of the simulations, there is a likely stability of the studied system. Although we have not performed classical MD of Apo Imp $\alpha$ , Takeda et al. [63] reinforce the instability of an N-terminally truncated Imp $\alpha$  in the absence of an NLS peptide. One way of computationally evaluating the N-terminally truncated Imp $\alpha$  would be applying long MDs to observe the effects of high protein flexibility. However, with a simple NM calculation, we could observe in general, based on the restraint energy profile, RMSF values and geometrical analysis, the higher flexibility of Imp $\alpha$  in the absence of the NLS.

The MD data with the Imp $\alpha$ -NplNLS complex not only supported the maintenance of the protein-peptide complex but also showed some of the major movements and interactions occurring at the complex interface. NMs promoted an analytical method to access the dynamics of the system, allowing the possibility of the recognition of new interactions and dribbling the convergence and conformational restrictions from standard MD. The main motions described here in NMs were also obtained in the PC analysis, in agreement with the comparison between the lowest NMs and the first PCs obtained from MD simulations in the lysozyme model [62] and a subunit of the GroEl chaperone [69]. The highlighted movements are of high occurrence in the protein's lifetime and are likely to be functionally important. Moreover, the low computational cost of NM is once more an attractive feature for application in biological systems.

## Conclusions

Computational approaches of MD and NM analysis were combined to evaluate the main of motions of Imp $\alpha$  and its interaction to a cNLS peptide. The bending motion may be involved in the NLS entrance and the accommodation of cargo protein depending on its size, whereas the twist motions may be involved in the NLS recognition and accommodation into the Imp $\alpha$  binding sites. The combination of these movements could allow local adjustments between the ARM repeats, which could improve the overall affinity to the cNLS. The absence of an NLS was also evaluated and may imply in a wider curvature of Apo Imp $\alpha$ , allowing for the IBB domain or cargo protein binding. Moreover, a higher twist in the middle of Apo Imp $\alpha$  was detected possibly due to the lack of interactions that stabilize and provide binding specificity, which could explain the challenges in crystallizing N-terminally truncated Apo Imp $\alpha$ .

The evaluation of salt bridges, hydrogen bonds and hydrophobic interactions corroborates the fundamental interactions between Imp $\alpha$  and NplNLS and gives additional support for interactions outside the classical binding pockets that are important during this process. The linker contacts in cNLS assist the adjustment of the peptide backbone, which helps the interactions between cNLS side chains and residues from major and minor binding grooves. In conclusion, MD simulations combined to NM analysis supported the maintenance of the Imp $\alpha$ -NplNLS complex exploring the conformational space and accessing the dynamics of the system with a lower computational cost. This approach may help to understand the affinities between Imp $\alpha$  and cNLS peptides and non-classic NLSs.

## Supporting Information

**S1 Movie. Movie of NM7.** Imp $\alpha$  is shown in a cartoon diagram, and it is colored according to each ARM repeat, from the N-terminal (blue) to the C-terminal (red) ends of the protein. The

NplNLS is represented in a cyan cartoon diagram positioned in an antiparallel configuration compared to Imp $\alpha$ .

(MP4)

**S2 Movie. Movie of NM8.** Imp $\alpha$  is shown in a cartoon diagram, and it is colored according to each ARM repeat, from the N-terminal (blue) to the C-terminal (red) ends of the protein. The NplNLS is represented in a cyan cartoon diagram positioned in an antiparallel configuration compared to Imp $\alpha$ .

(MP4)

**S3 Movie. Movie of NM9.** Imp $\alpha$  is shown in a cartoon diagram, and it is colored according to each ARM repeat, from the N-terminal (blue) to the C-terminal (red) ends of the protein. The NplNLS is represented in a cyan cartoon diagram positioned in an antiparallel configuration compared to Imp $\alpha$ .

(MP4)

**S1 Fig. Interactions from the starting structure.** Scheme of interactions of the starting structure for MD simulations. The representation is similar to the description of [Fig 7](#).

(PDF)

**S2 Fig. Convergence analysis.** Bar plot indicating the frequency of structures clustered within each reference (67,730 ps, 207,080 ps and 274,970 ps) from the MD simulations.

(PDF)

**S3 Fig. Structural alignment of NplNLS.** Structural representation of the NplNLSs from reference structures 67,730 ps (orange), 207,080 ps (green) and 274,970 ps (blue) aligned with the NplNLS from X-ray (red). (A) Stick diagram of the main chain, with major and minor sites indicated. (B-C) Stick diagram including side chains and the residues from each site.

(PDF)

**S4 Fig. Heatmap of the total restraint energy values from Imp $\alpha$  –NplNLS.** The values are from the structures generated from the NM-displacement of the most representative reference structure (207,080 ps) according to the convergence analysis. The X-axis is the displacement range, represented as values of mass-weighted root mean square (MRMS), and the Y-axis is the NM numbers. Lower values of energy (blue tons) indicate favorable conformations.

(PDF)

**S5 Fig. Heatmap of the total restraint energy values from Apo Imp $\alpha$ .** The values are from the structures generated from the NM-displacement. The X-axis is the displacement range, represented as values of MRMS, and the Y-axis is the NM numbers. Lower values of energy (blue tons) indicate favorable conformations.

(PDF)

**S6 Fig. Conformational exploration.** Box-plot of the RMSD distribution from the trajectories of standard MD (MD-1, MD-2 and MD-3) and NM-displacement (67,730 ps, 207,080 ps and 274,970 ps).

(PDF)

**S7 Fig. Residue fluctuations of Imp $\alpha$  –NplNLS.** Residue fluctuations based on C $\alpha$  RMSF of Imp $\alpha$  from the ensemble trajectories of standard MD (red) and NM-displacement (cyan).

(PDF)

**S8 Fig. Residue fluctuations of NplNLS.** Residue fluctuations based on  $C\alpha$  RMSF of NplNLS from the ensemble trajectories of standard MD (red) and NM-displacement (cyan).  
(PDF)

**S9 Fig. Residue fluctuations of apo  $\text{Imp}\alpha\Delta\text{IBB}$ .** Residue fluctuations based on  $C\alpha$  RMSF of  $\text{Imp}\alpha$  from the trajectories of the NM-displacement of  $\text{Imp}\alpha$ –NplNLS (from reference structure 207,080 ps; cyan) and Apo  $\text{Imp}\alpha\Delta\text{IBB}$  (black).  
(PDF)

**S10 Fig. Collectivity from  $\text{Imp}\alpha$ –NplNLS.** The collectivity values are plotted for each NM, and a smoothed line is fitted (blue line), representing the data tendency. The shaded area is the confidence interval around the smoothed line calculated with the ggplot package (<http://ggplot2.org/>) in R.  
(PDF)

**S11 Fig. Main motions observed in NMs10–13, NM17 and NM18.** The vector arrows indicating the motions are shown. The  $\text{Imp}\alpha$  is displayed in a cartoon diagram, with each ARM colored from blue to red, corresponding to N to C-terminals. The NplNLS (cyan) is in a cartoon representation and is positioned in an antiparallel configuration compared to  $\text{Imp}\alpha$ .  
(PDF)

**S12 Fig. Main motions observed in PCs1–3 from standard MD.** The vector arrows indicating the motions are shown. The  $\text{Imp}\alpha$  is displayed in a  $C\alpha$  representation, with each ARM colored from blue to red, corresponding to N to C-terminals. The NplNLS (cyan) is in  $C\alpha$  representation and is positioned in an antiparallel configuration compared to  $\text{Imp}\alpha$ .  
(PDF)

**S13 Fig. Angles between helices of  $\text{Imp}\alpha$ –NplNLS for NMs10–13, NM17 and NM18.** The ARM groups considered for each angle calculation are depicted with different color assignments, similar to Fig 5.  
(PDF)

**S1 Table. Backbone RMSD values from structural alignment.** The reference structures are aligned with the  $\text{Imp}\alpha$ –NplNLS crystallographic structure (PDB ID 3UL1).  
(PDF)

**S2 Table. Salt bridges occupancies.** The occupancies of salt bridges between NplNLS and  $\text{Imp}\alpha$  in standard MD and NM-displacement. Interactions that were above 50% of occupancy are highlighted in gray.  
(PDF)

**S3 Table. Hydrogen bonds occupancies.** The occupancies of hydrogen bonds between NplNLS (blue) and  $\text{Imp}\alpha$  (green) in standard MD and NM-displacement. Interactions that were above 50% of occupancy are highlighted in gray.  
(PDF)

**S4 Table. Hydrophobic contacts occupancies.** The occupancies of hydrophobic contacts between NplNLS and  $\text{Imp}\alpha$  in standard MD and NM-displacement. Interactions that were above 50% of occupancy are highlighted in gray.  
(PDF)

**S5 Table. Radius of curvature and average angles for the crystallographic structures.** PDB IDs: 1pjm, 1pjn, 3tpm, 3ukw, 3ukx, 3uky, 3ukz, 3ul1 and 3uvu.  
(PDF)

## Acknowledgments

This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) [grant numbers 2012/19447-2; 2014/21976-9] and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [grant number 142110/2012-4]. We are grateful to Prof. Dr. Cesar Martins for providing computational access for running part of the simulations.

## Author Contributions

Conceived and designed the experiments: MTG AAST ASKB NL. Performed the experiments: MTG. Analyzed the data: MTG AAST ASKB NL. Wrote the paper: MTG AAST.

## References

1. Kobe B, Kajava AV. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends in biochemical sciences*. 2000; 25(10):509–515. doi: [10.1016/S0968-0004\(00\)01667-4](https://doi.org/10.1016/S0968-0004(00)01667-4) PMID: [11050437](https://pubmed.ncbi.nlm.nih.gov/11050437/)
2. Kappel C, Zachariae U, Dölker N, Grubmüller H. An unusual hydrophobic core confers extreme flexibility to HEAT repeat proteins. *Biophysical journal*. 2010; 99(5):1596–1603. doi: [10.1016/j.bpj.2010.06.032](https://doi.org/10.1016/j.bpj.2010.06.032) PMID: [20816072](https://pubmed.ncbi.nlm.nih.gov/20816072/)
3. Doyle L, Hallinan J, Bolduc J, Parmeggiani F, Baker D, Stoddard BL, et al. Rational design of  $\alpha$ -helical tandem repeat proteins with closed architectures. *Nature*. 2015; 528(7583):585–588. doi: [10.1038/nature16191](https://doi.org/10.1038/nature16191) PMID: [26675735](https://pubmed.ncbi.nlm.nih.gov/26675735/)
4. Settanni G, Serquera D, Marszalek PE, Paci E, Itzhaki LS. Effects of ligand binding on the mechanical properties of ankyrin repeat protein gankyrin. *PLoS Comput Biol*. 2013; 9(1):e1002864. doi: [10.1371/journal.pcbi.1002864](https://doi.org/10.1371/journal.pcbi.1002864) PMID: [23341763](https://pubmed.ncbi.nlm.nih.gov/23341763/)
5. Christie M, Chang CW, Róna G, Smith KM, Stewart AG, Takeda AA, et al. Structural biology and regulation of protein import into the nucleus. *Journal of molecular biology*. 2015. PMID: [26523678](https://pubmed.ncbi.nlm.nih.gov/26523678/)
6. Conti E, Uy M, Leighton L, Blobel G, Kuriyan J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin  $\alpha$ . *Cell*. 1998; 94(2):193–204. doi: [10.1016/S0092-8674\(00\)81419-1](https://doi.org/10.1016/S0092-8674(00)81419-1) PMID: [9695948](https://pubmed.ncbi.nlm.nih.gov/9695948/)
7. Radu A, Blobel G, Moore MS. Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins. *Proceedings of the National Academy of Sciences*. 1995; 92(5):1769–1773. doi: [10.1073/pnas.92.5.1769](https://doi.org/10.1073/pnas.92.5.1769)
8. Cingolani G, Petosa C, Weis K, Müller CW. Structure of importin- $\beta$  bound to the IBB domain of importin- $\alpha$ . *Nature*. 1999; 399(6733):221–229. doi: [10.1038/20367](https://doi.org/10.1038/20367) PMID: [10353244](https://pubmed.ncbi.nlm.nih.gov/10353244/)
9. Görlich D, Kostka S, Kraft R, Dingwall C, Laskey RA, Hartmann E, et al. Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Current Biology*. 1995; 5(4):383–392. doi: [10.1016/S0960-9822\(95\)00079-0](https://doi.org/10.1016/S0960-9822(95)00079-0) PMID: [7627554](https://pubmed.ncbi.nlm.nih.gov/7627554/)
10. Görlich D, Henklein P, Laskey R, Hartmann E. A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus. *The EMBO Journal*. 1996; 15(8):1810. PMID: [8617226](https://pubmed.ncbi.nlm.nih.gov/8617226/)
11. Moroianu J, Blobel G, Radu A. Nuclear protein import: Ran-GTP dissociates the karyopherin alpha-beta heterodimer by displacing alpha from an overlapping binding site on beta. *Proceedings of the National Academy of Sciences*. 1996; 93(14):7059–7062. doi: [10.1073/pnas.93.14.7059](https://doi.org/10.1073/pnas.93.14.7059)
12. Lee SJ, Matsuura Y, Liu SM, Stewart M. Structural basis for nuclear import complex dissociation by RanGTP. *Nature*. 2005; 435(7042):693–696. doi: [10.1038/nature03578](https://doi.org/10.1038/nature03578) PMID: [15864302](https://pubmed.ncbi.nlm.nih.gov/15864302/)
13. Matsuura Y, Stewart M. Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. *The EMBO journal*. 2005; 24(21):3681–3689. doi: [10.1038/sj.emboj.7600843](https://doi.org/10.1038/sj.emboj.7600843) PMID: [16222336](https://pubmed.ncbi.nlm.nih.gov/16222336/)
14. Matsuura Y, Stewart M. Structural basis for the assembly of a nuclear export complex. *Nature*. 2004; 432(7019):872–877. doi: [10.1038/nature03144](https://doi.org/10.1038/nature03144) PMID: [15602554](https://pubmed.ncbi.nlm.nih.gov/15602554/)
15. Riggleman B, Wieschaus E, Schedl P. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. *Genes & development*. 1989; 3(1):96–113. doi: [10.1101/gad.3.1.96](https://doi.org/10.1101/gad.3.1.96)
16. Kobe B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin  $\alpha$ . *Nature Structural & Molecular Biology*. 1999; 6(4):388–397. doi: [10.1038/7625](https://doi.org/10.1038/7625)



17. Marfori M, Lonhienne TG, Forwood JK, Kobe B. Structural Basis of High-Affinity Nuclear Localization Signal Interactions with Importin- $\alpha$ . *Traffic*. 2012; 13(4):532–548. doi: [10.1111/j.1600-0854.2012.01329.x](https://doi.org/10.1111/j.1600-0854.2012.01329.x) PMID: [22248489](https://pubmed.ncbi.nlm.nih.gov/22248489/)
18. Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of cell biology*. 1988; 107(3):841–849. doi: [10.1083/jcb.107.3.841](https://doi.org/10.1083/jcb.107.3.841) PMID: [3417784](https://pubmed.ncbi.nlm.nih.gov/3417784/)
19. Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. *Cell*. 1984; 39(3):499–509. doi: [10.1016/0092-8674\(84\)90457-4](https://doi.org/10.1016/0092-8674(84)90457-4) PMID: [6096007](https://pubmed.ncbi.nlm.nih.gov/6096007/)
20. Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell*. 1991; 64(3):615–623. doi: [10.1016/0092-8674\(91\)90245-T](https://doi.org/10.1016/0092-8674(91)90245-T) PMID: [1991323](https://pubmed.ncbi.nlm.nih.gov/1991323/)
21. Fontes MR, Teh T, Jans D, Brinkworth RI, Kobe B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin- $\alpha$ . *Journal of Biological Chemistry*. 2003; 278(30):27981–27987. doi: [10.1074/jbc.M303275200](https://doi.org/10.1074/jbc.M303275200) PMID: [12695505](https://pubmed.ncbi.nlm.nih.gov/12695505/)
22. Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, et al. Six classes of nuclear localization signals specific to different binding grooves of importin  $\alpha$ . *Journal of Biological Chemistry*. 2009; 284(1):478–485. doi: [10.1074/jbc.M807017200](https://doi.org/10.1074/jbc.M807017200) PMID: [19001369](https://pubmed.ncbi.nlm.nih.gov/19001369/)
23. Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *Journal of Biological Chemistry*. 2001; 276(2):1317–1325. doi: [10.1074/jbc.M008522200](https://doi.org/10.1074/jbc.M008522200) PMID: [11038364](https://pubmed.ncbi.nlm.nih.gov/11038364/)
24. Fontes MR, Teh T, Kobe B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin- $\alpha$ . *Journal of molecular biology*. 2000; 297(5):1183–1194. doi: [10.1006/jmbi.2000.3642](https://doi.org/10.1006/jmbi.2000.3642) PMID: [10764582](https://pubmed.ncbi.nlm.nih.gov/10764582/)
25. Conti E, Kuriyan J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin  $\alpha$ . *Structure*. 2000; 8(3):329–338. doi: [10.1016/S0969-2126\(00\)00107-6](https://doi.org/10.1016/S0969-2126(00)00107-6) PMID: [10745017](https://pubmed.ncbi.nlm.nih.gov/10745017/)
26. Chen MH, Ben-Efraim I, Mitrousis G, Walker-Kopp N, Sims PJ, Cingolani G. Phospholipid scramblase 1 contains a nonclassical nuclear localization signal with unique binding site in importin  $\alpha$ . *Journal of Biological Chemistry*. 2005; 280(11):10599–10606. doi: [10.1074/jbc.M413194200](https://doi.org/10.1074/jbc.M413194200) PMID: [15611084](https://pubmed.ncbi.nlm.nih.gov/15611084/)
27. Cutress ML, Whitaker HC, Mills IG, Stewart M, Neal DE. Structural basis for the nuclear import of the human androgen receptor. *Journal of cell science*. 2008; 121(7):957–968. doi: [10.1242/jcs.022103](https://doi.org/10.1242/jcs.022103) PMID: [18319300](https://pubmed.ncbi.nlm.nih.gov/18319300/)
28. Giesecke A, Stewart M. Novel binding of the mitotic regulator TPX2 (target protein for Xenopus kinesin-like protein 2) to importin- $\alpha$ . *Journal of Biological Chemistry*. 2010; 285(23):17628–17635. doi: [10.1074/jbc.M110.102343](https://doi.org/10.1074/jbc.M110.102343) PMID: [20335181](https://pubmed.ncbi.nlm.nih.gov/20335181/)
29. Yang SN, Takeda AA, Fontes MR, Harris JM, Jans DA, Kobe B. Probing the specificity of binding to the major nuclear localization sequence-binding site of importin- $\alpha$  using oriented peptide library screening. *Journal of Biological Chemistry*. 2010; 285(26):19935–19946. doi: [10.1074/jbc.M109.079574](https://doi.org/10.1074/jbc.M109.079574) PMID: [20406804](https://pubmed.ncbi.nlm.nih.gov/20406804/)
30. de Barros AC, Takeda AA, Chang CW, Kobe B, Fontes MR. Structural basis of nuclear import of flap endonuclease 1 (FEN1). *Acta Crystallographica Section D: Biological Crystallography*. 2012; 68(7):743–750. doi: [10.1107/S0907444912010281](https://doi.org/10.1107/S0907444912010281)
31. Forwood JK, Lange A, Zachariae U, Marfori M, Preast C, Grubmüller H, et al. Quantitative structural analysis of importin- $\beta$  flexibility: paradigm for solenoid protein structures. *Structure*. 2010; 18(9):1171–1183. doi: [10.1016/j.str.2010.06.015](https://doi.org/10.1016/j.str.2010.06.015) PMID: [20826343](https://pubmed.ncbi.nlm.nih.gov/20826343/)
32. Makkerh JP, Dingwall C, Laskey RA. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Current Biology*. 1996; 6(8):1025–1027. doi: [10.1016/S0960-9822\(02\)00648-6](https://doi.org/10.1016/S0960-9822(02)00648-6) PMID: [8805337](https://pubmed.ncbi.nlm.nih.gov/8805337/)
33. Dingwall C, Laskey R. Nucleoplasmin: the archetypal molecular chaperone. In: *Seminars in cell biology*. vol. 1; 1990. p. 11–17.
34. Eswar N, Webb B, Marti-Renom MA, Madhusudhan M, Eramian D, Shen My, et al. Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics*. 2006;p. 5–6.
35. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*. 2009; 66(1):12–21. doi: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073)
36. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology*. 2007; 372(3):774–797. doi: [10.1016/j.jmb.2007.05.022](https://doi.org/10.1016/j.jmb.2007.05.022) PMID: [17681537](https://pubmed.ncbi.nlm.nih.gov/17681537/)
37. Hess B. P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*. 2008; 4(1):116–122. doi: [10.1021/ct700200b](https://doi.org/10.1021/ct700200b) PMID: [26619985](https://pubmed.ncbi.nlm.nih.gov/26619985/)

38. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry*. 2013; 34(25):2135–2145. doi: [10.1002/jcc.23354](https://doi.org/10.1002/jcc.23354) PMID: [23832629](https://pubmed.ncbi.nlm.nih.gov/23832629/)
39. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*. 1983; 79(2):926–935. doi: [10.1063/1.445869](https://doi.org/10.1063/1.445869)
40. MacKerell AD, Bashford D, Bellott M, Dunbrack R, Evanseck J, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*. 1998; 102(18):3586–3616. doi: [10.1021/jp973084f](https://doi.org/10.1021/jp973084f) PMID: [24889800](https://pubmed.ncbi.nlm.nih.gov/24889800/)
41. Hess B, Bekker H, Berendsen HJ, Fraaije JG, et al. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry*. 1997; 18(12):1463–1472. doi: [10.1002/\(SICI\)1096-987X\(199709\)18:12%3C1463::AID-JCC4%3E3.3.CO;2-L](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%3C1463::AID-JCC4%3E3.3.CO;2-L)
42. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *The Journal of chemical physics*. 1995; 103(19):8577–8593. doi: [10.1063/1.470117](https://doi.org/10.1063/1.470117)
43. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *The Journal of chemical physics*. 2007; 126(1):014101. doi: [10.1063/1.2408420](https://doi.org/10.1063/1.2408420) PMID: [17212484](https://pubmed.ncbi.nlm.nih.gov/17212484/)
44. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*. 1981; 52(12):7182–7190. doi: [10.1063/1.328693](https://doi.org/10.1063/1.328693)
45. Lyman E, Zuckerman DM. Ensemble-based convergence analysis of biomolecular trajectories. *Biophysical journal*. 2006; 91(1):164–172. doi: [10.1529/biophysj.106.082941](https://doi.org/10.1529/biophysj.106.082941) PMID: [16617086](https://pubmed.ncbi.nlm.nih.gov/16617086/)
46. Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. *Journal of computational chemistry*. 2009; 30(10):1545–1614. doi: [10.1002/jcc.21287](https://doi.org/10.1002/jcc.21287) PMID: [19444816](https://pubmed.ncbi.nlm.nih.gov/19444816/)
47. Floquet N, Costa MG, Batista PR, Renault P, Bisch PM, Raussin F, et al. Conformational Equilibrium of CDK/Cyclin Complexes by Molecular Dynamics with Excited Normal Modes. *Biophysical journal*. 2015; 109(6):1179–1189. doi: [10.1016/j.bpj.2015.07.003](https://doi.org/10.1016/j.bpj.2015.07.003) PMID: [26255588](https://pubmed.ncbi.nlm.nih.gov/26255588/)
48. Brüschweiler R. Collective protein dynamics and nuclear spin relaxation. *The Journal of chemical physics*. 1995; 102(8):3396–3403. doi: [10.1063/1.469213](https://doi.org/10.1063/1.469213)
49. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. *Protein engineering*. 2001; 14(1):1–6. doi: [10.1093/protein/14.1.1](https://doi.org/10.1093/protein/14.1.1) PMID: [11287673](https://pubmed.ncbi.nlm.nih.gov/11287673/)
50. Bakan A, Meireles LM, Bahar I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics*. 2011; 27(11):1575–1577. doi: [10.1093/bioinformatics/btr168](https://doi.org/10.1093/bioinformatics/btr168) PMID: [21471012](https://pubmed.ncbi.nlm.nih.gov/21471012/)
51. Seeber M, Cecchini M, Rao F, Settanni G, Caffisch A. Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics*. 2007; 23(19):2625–2627. doi: [10.1093/bioinformatics/btm378](https://doi.org/10.1093/bioinformatics/btm378) PMID: [17717034](https://pubmed.ncbi.nlm.nih.gov/17717034/)
52. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics*. 1996; 14(1):33–38. doi: [10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: [8744570](https://pubmed.ncbi.nlm.nih.gov/8744570/)
53. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*. 2010; 66(4):486–501. doi: [10.1107/S0907444910007493](https://doi.org/10.1107/S0907444910007493)
54. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *Journal of computational and graphical statistics*. 1996; 5(3):299–314. doi: [10.2307/1390807](https://doi.org/10.2307/1390807)
55. Schrödinger L. The PyMOL Molecular Graphics System, Version 1.3 r1; Schrödinger, LLC. New York. 2010.
56. Keskin O, Durell SR, Bahar I, Jernigan RL, Covell DG. Relating molecular flexibility to function: a case study of tubulin. *Biophysical journal*. 2002; 83(2):663–680. doi: [10.1016/S0006-3495\(02\)75199-0](https://doi.org/10.1016/S0006-3495(02)75199-0) PMID: [12124255](https://pubmed.ncbi.nlm.nih.gov/12124255/)
57. Kajava AV. Review: proteins with repeated sequence—structural prediction and modeling. *Journal of structural biology*. 2001; 134(2):132–144. doi: [10.1006/jsbi.2000.4328](https://doi.org/10.1006/jsbi.2000.4328) PMID: [11551175](https://pubmed.ncbi.nlm.nih.gov/11551175/)
58. Lee SJ, Imamoto N, Sakai H, Nakagawa A, Kose S, Koike M, et al. The adoption of a twisted structure of importin- $\beta$  is essential for the protein-protein interaction required for nuclear transport. *Journal of molecular biology*. 2000; 302(1):251–264. doi: [10.1006/jmbi.2000.4055](https://doi.org/10.1006/jmbi.2000.4055) PMID: [10964573](https://pubmed.ncbi.nlm.nih.gov/10964573/)
59. Brooks B, Karplus M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proceedings of the National Academy of Sciences*. 1985; 82(15):4995–4999. doi: [10.1073/pnas.82.15.4995](https://doi.org/10.1073/pnas.82.15.4995)
60. Ichiye T, Karplus M. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Bioinformatics*. 1991; 11(3):205–217. doi: [10.1002/prot.340110305](https://doi.org/10.1002/prot.340110305)

61. Pumroy RA, Ke S, Hart DJ, Zachariae U, Cingolani G. Molecular determinants for nuclear import of influenza A PB2 by importin  $\alpha$  isoforms 3 and 7. *Structure*. 2015; 23(2):374–384. doi: [10.1016/j.str.2014.11.015](https://doi.org/10.1016/j.str.2014.11.015) PMID: [25599645](https://pubmed.ncbi.nlm.nih.gov/25599645/)
62. Hayward S, Kitao A, Berendsen HJ. Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins: Structure, Function, and Bioinformatics*. 1997; 27(3):425–437. doi: [10.1002/\(SICI\)1097-0134\(199703\)27:3%3C425::AID-PROT10%3E3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0134(199703)27:3%3C425::AID-PROT10%3E3.0.CO;2-N)
63. Takeda AA, Freitas FZ, Magro AJ, Bernardes NE, Fernandes CA, Goncalves RD, et al. Biophysical Characterization of the Recombinant Importin- $\alpha$  from *Neurospora crassa*. *Protein and peptide letters*. 2013; 20(1):8–16. doi: [10.2174/092986613804096829](https://doi.org/10.2174/092986613804096829) PMID: [22789101](https://pubmed.ncbi.nlm.nih.gov/22789101/)
64. Falces J, Arregi I, Konarev PV, Urbaneja MA, Svergun DI, Taneva SG, et al. Recognition of nucleoplasmin by its nuclear transport receptor importin  $\alpha/\beta$ : insights into a complete import complex. *Biochemistry*. 2010; 49(45):9756–9769. doi: [10.1021/bi101179g](https://doi.org/10.1021/bi101179g) PMID: [20925424](https://pubmed.ncbi.nlm.nih.gov/20925424/)
65. Kirby TW, Gassman NR, Smith CE, Pedersen LC, Gabel SA, Sobhany M, et al. Nuclear Localization of the DNA Repair Scaffold XRCC1: Uncovering the Functional Role of a Bipartite NLS. *Scientific reports*. 2015; 5. doi: [10.1038/srep13405](https://doi.org/10.1038/srep13405)
66. de Barros AC, Takeda AA, Dreyer TR, Velazquez-Campoy A, Kobe B, Fontes MR. Structural and calorimetric studies demonstrate that Xeroderma pigmentosum type G (XPG) can be imported to the nucleus by a classical nuclear import pathway via a monopartite NLS sequence. *Journal of Molecular Biology*. 2016.
67. Bernardes NE, Takeda AA, Dreyer TR, Freitas FZ, Bertolini MC, Fontes MR. Structure of Importin- $\alpha$  from a Filamentous Fungus in Complex with a Classical Nuclear Localization Signal. *PloS one*. 2015; 10(6):e0128687. doi: [10.1371/journal.pone.0128687](https://doi.org/10.1371/journal.pone.0128687) PMID: [26091498](https://pubmed.ncbi.nlm.nih.gov/26091498/)
68. Alexandrov V, Lehnert U, Echols N, Milburn D, Engelman D, Gerstein M. Normal modes for predicting protein motions: a comprehensive database assessment and associated Web tool. *Protein science*. 2005; 14(3):633–643. doi: [10.1110/ps.04882105](https://doi.org/10.1110/ps.04882105) PMID: [15722444](https://pubmed.ncbi.nlm.nih.gov/15722444/)
69. Skjaerven L, Martinez A, Reuter N. Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit. *Proteins: Structure, Function, and Bioinformatics*. 2011; 79(1):232–243. doi: [10.1002/prot.22875](https://doi.org/10.1002/prot.22875)
70. Grossfield A, Zuckerman DM. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annual reports in computational chemistry*. 2009; 5:23–48. doi: [10.1016/S1574-1400\(09\)00502-7](https://doi.org/10.1016/S1574-1400(09)00502-7) PMID: [20454547](https://pubmed.ncbi.nlm.nih.gov/20454547/)
71. Genheden S, Ryde U. Will molecular dynamics simulations of proteins ever reach equilibrium? *Physical Chemistry Chemical Physics*. 2012; 14(24):8662–8677. PMID: [22614001](https://pubmed.ncbi.nlm.nih.gov/22614001/)