Mechanism of Action of Cyclophilin A Explored by Metadynamics Simulations

Vanessa Leone¹,², Gianluca Lattanzi³,⁴, Carla Molteni⁵, Paolo Carloni¹,²

¹ International School for Advanced Studies (SISSA), Trieste, Italy, ² IIT—Italian Institute of Technology and DEMOCRITOS, Trieste, Italy, ³ University of Bari, Bari, Italy, ⁴ TRES and INFN, Bari, Italy, ⁵ Department of Physics, King’s College, London, United Kingdom

Abstract

Trans/cis prolyl isomerisation is involved in several biological processes, including the development of numerous diseases. In the HIV-1 capsid protein (CA), such a process takes place in the uncoating and recruitment of the virion and is catalyzed by cyclophilin A (CypA). Here, we use metadynamics simulations to investigate the isomerization of CA’s model substrate HAGPIA in water and in its target protein CypA. Our results allow us to propose a novel mechanistic hypothesis, which is finally consistent with all of the available molecular biology data.

Introduction

Proline trans/cis isomerization takes part in several fundamental biological processes, including protein folding [1–3], immune response [3,4], ion channel gating [5] and cellular signalling [3,4,6]. The process, which is also associated to the development of a variety of diseases including HIV-1 infection [7,8], carcinogenesis [6] and Alzheimer’s [9], is catalyzed by prolyl isomerase enzymes [1,10,11].

The best characterized isomerization in vivo and in vitro occurs in the uncoating and recruitment of the human HIV-1 capsid (CA) protein in the virions [7], and it is catalyzed by cyclophilin A isomerase (CypA) [12]. At the structural level, CypA features α-helices flanking a beta-barrel, while CA is made of several α-helices connected by loops (Figure 1). In the X-ray structure of the complex [13], the targeted proline-containing backbone unit (G89-P90), is accommodated in a hydrophobic pocket of CypA that stabilizes in the proposed mechanisms are important for the stabilization in the enzyme active site by a network of hydrogen bonds [15,22,23].

In spite of the great progress in describing the catalytic process, key mechanistic issues need to be addressed. Kinetic studies [17] suggested that the overall process involves trans and cis forms in solution and in complex with the protein. However, the studies so far consider mostly TS stabilization. Most importantly, the current proposed mechanism cannot explain a plethora of molecular biology data. These studies show that residues not involved in TS stabilization in the proposed mechanisms are important for the function, as their mutations cause a decrease of kcat/Km. Indeed, kcat/Km passes from 1.6 10⁷ M⁻¹ s⁻¹ in the wild type to 1.6 10⁶ M⁻¹ s⁻¹ by mutating the fully conserved R55 residue with Ala [20]. Although this mutation was proposed originally to destroy an H-bond stabilizing uniquely the TS [15,24], such H-bond was subsequently ruled out in computational works...
Mechanism of Action of Cyclophilin A

**Author Summary**

Peptidyl prolyl isomerases are ubiquitous enzymes whose actions are crucial in several biological processes, such as, for instance, in cellular signalling and in the onset of several diseases, e.g., HIV infection. Therefore, these isomerases are promising targets for the design of new drugs. For this purpose, we need to understand their molecular mechanism of action. One of the most characterized peptidyl prolyl isomerases is cyclophilin A. Previous studies characterized the roles of several protein regions in isomerase function. However, there are still experimentally identified important portions of the protein whose specific actions in the mechanism are still not known. Here, we address this problem by an extensive computational study of cyclophilin A and a substrate peptide that is part of the HIV-1 capsid protein. We present a novel four-step mechanism of the whole enzymatic process, which is consistent with all of the available experimental data. Moreover, these steps can be used as targets for the development of drugs, e.g., for HIV-1 infection.

[14,15,18] and so far no functional role has been ascribed to R55. In addition, the H54Q mutation, along with the I57V polymorphism, causes a decrease of k_{cat}/K_{m}, although these residues are by far the most stable in aqueous solution. In addition, the most likely barriers are only compared at the qualitative level to discriminate populations of the equilibrium conformations, whilst the calculated energy differences between minima are used to predict the relative influence on the barriers is more significant. Therefore, here the free energy calculations. Based on this comparison, we find that the force-field choice allows a very efficient sampling because of its relatively small computational cost. In spite of its limitations, [30] this force field is available experimental data. Moreover, these steps can be used as targets for the development of drugs, e.g., for HIV-1 infection.

Free Energy Calculations of Proline cis/trans Isomerisation

Free energies were calculated using the metadynamics method in its bias exchange variant [26,27]. This approach consists in a continuous addition of a history dependent potential energy that enforces the dynamics to explore conformations that were not previously visited. Briefly, the forces acting on each atom, or centroid, are corrected by a history dependent contribution, obtained as the derivative of the history dependent potential energy V_{\tau} with respect to the atom coordinate. V_{\tau} is given by the sum of a set of Gaussian functions centered on the values s_{i} of each chosen collective variable (CV) s as time t:

\[ V_{\tau} = \sum_{i < t} w \exp \left( -\frac{(s - s_{i})^{2}}{2\sigma^{2}} \right) \] (1)

The time interval between the addition of two Gaussian functions \( \tau \), as well as the Gaussian height \( w \) and Gaussian width \( \sigma \), are tuned to optimize the ratio between accuracy and computational cost (Table S2). Eventually, after exploring all conformations defined within the CV space, the probability distribution of Gaussians becomes flat, and the free energy profile does not change any more. At this stage, the free energy surface can be easily reconstructed as the opposite of the sum of all Gaussians.

Here we use the bias exchange variant of this method [27], in which several trajectories x^a, x^b with different history dependent potentials (for instance two CVS \( a \) and \( b \)), \( V_{\tau}^a, \ V_{\tau}^b \), are run in parallel. At specific time intervals (of the order of few ps), \( x^a \) and \( x^b \) are swapped with a probability \( P_{ab} \) evaluated by the standard Metropolis scheme [39].

\[ P_{ab} = \begin{cases} 1 & \text{if } \Delta \leq 0 \\ e^{-\Delta} & \text{if } \Delta > 0 \end{cases} \] (2)

\[ \Delta = \beta \left( V_{\tau}^a (x^b, t) + V_{\tau}^b (x^a, t) - V_{\tau}^a (x^a, t) - V_{\tau}^b (x^b, t) \right) \]

The initial structural models of HAGPIA/ in the free state (PEPT-WAT) and in complex with CypA (PEPT-CypA) were obtained from the CA N-terminal domain/CypA X-ray structure at 2.0 \( \text{Å} \) resolution (PDBID:1M9C) [13]. The X-ray data were collected at 100 K and slightly basic pH 3.0 [13]. The peptide and the protein were considered in their zwitterionic form. The protonation states of all histidine residues were determined by pK_a calculations based on the H++ server [31,32] and by visual inspection of the hydrogen bond network. All His residues were protonated at N81. The result for PEPT-WAT is consistent with the NMR study on CA N-terminal domain at pH 7 [33].

PEPT-WAT and PEPT-CypA were solvated in a 32 \( \text{Å} \times 32 \times 32 \) boxes, containing 1,423 water molecules (for a total of 4,349 atoms) and 8,233 water molecules and a Cl^− ion (for a total of 27,283 atoms), respectively. The ion was added to neutralize the system. The Amber99 all atom force field [29] was used for the protein and the chloride ion; the TIP3P model [34] was used for water. After minimization up to a convergence of \( 10^{-4} \) kcal/mol (conjugate gradient algorithm), the system was equilibrated for 1 ns (time step of 1 fs) in an isothermal-isobaric ensemble (1 atm, 310 K) with the Langevin barostat [35] (the oscillation period and the decay coefficient were set to 200 fs and 100 fs, respectively) and thermostat [36] (coupling coefficient 1 ps^{-1}). We use the Particle Mesh Ewald scheme [37] with 12 Å cutoff and 0.75 Å-spaced Fourier grid; we assume a dielectric constant of 1. Van der Walls interaction cutoff was set to 12 Å. Minimization and equilibration were performed with the NAMD 2.6 program suite [38].

**Methods**

**Structural Models and MD Calculations**

The initial structural models of HAGPIA/ in the free state (PEPT-WAT) and in complex with CypA (PEPT-CypA) were obtained from the CA N-terminal domain/CypA X-ray structure at 2.0 \( \text{Å} \) resolution (PDBID:1M9C) [13]. The X-ray data were collected at 100 K and slightly basic pH 3.0 [13]. The peptide and the protein were considered in their zwitterionic form. The protonation states of all histidine residues were determined by pK_a calculations based on the H++ server [31,32] and by visual inspection of the hydrogen bond network. All His residues were protonated at N81. The result for PEPT-WAT is consistent with the NMR study on CA N-terminal domain at pH 7 [33].

PEPT-WAT and PEPT-CypA were solvated in a 32 \( \text{Å} \times 32 \times 32 \) boxes, containing 1,423 water molecules (for a total of 4,349 atoms) and 8,233 water molecules and a Cl^− ion (for a total of 27,283 atoms), respectively. The ion was added to neutralize the system. The Amber99 all atom force field [29] was used for the protein and the chloride ion; the TIP3P model [34] was used for water. After minimization up to a convergence of \( 10^{-4} \) kcal/mol (conjugate gradient algorithm), the system was equilibrated for 1 ns (time step of 1 fs) in an isothermal-isobaric ensemble (1 atm, 310 K) with the Langevin barostat [35] (the oscillation period and the decay coefficient were set to 200 fs and 100 fs, respectively) and thermostat [36] (coupling coefficient 1 ps^{-1}). We use the Particle Mesh Ewald scheme [37] with 12 Å cutoff and 0.75 Å-spaced Fourier grid; we assume a dielectric constant of 1. Van der Walls interaction cutoff was set to 12 Å. Minimization and equilibration were performed with the NAMD 2.6 program suite [38].

Free Energy Calculations of Proline cis/trans Isomerisation

Free energies were calculated using the metadynamics method in its bias exchange variant [26,27]. This approach consists in a continuous addition of a history dependent potential energy that enforces the dynamics to explore conformations that were not previously visited. Briefly, the forces acting on each atom, or centroid, are corrected by a history dependent contribution, obtained as the derivative of the history dependent potential energy \( V_{\tau} \) with respect to the atom coordinate. \( V_{\tau} \) is given by the sum of a set of Gaussian functions centered on the values \( s_{i} \) of each chosen collective variable (CV) \( s \) as time \( t \):

\[ V_{\tau} = \sum_{i < t} w \exp \left( -\frac{(s - s_{i})^{2}}{2\sigma^{2}} \right) \] (1)

The time interval between the addition of two Gaussian functions \( \tau \), as well as the Gaussian height \( w \) and Gaussian width \( \sigma \), are tuned to optimize the ratio between accuracy and computational cost (Table S2). Eventually, after exploring all conformations defined within the CV space, the probability distribution of Gaussians becomes flat, and the free energy profile does not change any more. At this stage, the free energy surface can be easily reconstructed as the opposite of the sum of all Gaussians.

Here we use the bias exchange variant of this method [27], in which several trajectories \( x^{a}, x^{b} \) with different history dependent potentials (for instance two CVS \( a \) and \( b \)), \( V_{\tau}^{a}, V_{\tau}^{b} \), are run in parallel. At specific time intervals (of the order of few ps), \( x^{a} \) and \( x^{b} \) are swapped with a probability \( P_{ab} \) evaluated by the standard Metropolis scheme [39].

\[ P_{ab} = \begin{cases} 1 & \text{if } \Delta \leq 0 \\ e^{-\Delta} & \text{if } \Delta > 0 \end{cases} \] (2)

\[ \Delta = \beta \left( V_{\tau}^{a} (x^{b}, t) + V_{\tau}^{b} (x^{a}, t) - V_{\tau}^{a} (x^{a}, t) - V_{\tau}^{b} (x^{b}, t) \right) \]
Human CypA is a cytoplasmatic single peptide chain, 165 amino acids long. Top: The X-ray structure shows that it is a \(\beta\)-barrel with eight antiparallel \(\beta\)-strands and two \(\alpha\)-helices flanking the \(\beta\)-barrel [13]. These secondary structures are connected by several flexible loops. The \(\beta\)-barrel contains the active site for cis/trans prolyl isomerization. The CA protein is composed by two helices connected by a loop. This contains the G89-P90 target peptide bond (Bottom, right). Bottom, left: Transition state of the cis-trans isomerisation, emerging from this and previous [14,18,22] calculations.

doi:10.1371/journal.pcbi.1000309.g001
where \( k_B \) is the Boltzmann constant and \( T \) is the temperature. This method was shown to provide an efficient sampling of the conformational space [27].

The free energies associated with cis/trans isomerization of the G3P4 bond in PEPT-WAT and PEPT-CypA were calculated as a function of pairs of CVs in the canonical ensemble. The free energies were first calculated as a function of the \( \phi \) and \( \psi \) dihedral angles (Figure 2 and Table S2), because these angles have been shown to be essential to describe the isomerisation process [25]. The resulting free energy plot was used to define the minima and

| Source of Data | Isomerized Peptide | Thermodynamic Parameter | Values (kcal/mol) | Ref |
|----------------|-------------------|-------------------------|-------------------|-----|
| Experiment     | SUC-Ala-Phe-Pro-Phe-pNA<sup>b</sup> | \( \Delta G_{cis-trans} \) | 0.6 ± 0.2          | [17]|
|                | SUC-Ala-Phe-Pro-Phe-pNa<sup>b</sup> | \( \Delta G_{cis-trans} \) | 20.1 ± 12.9       | [17]|
|                | SUC-Ala-Pro-Phe-pNA<sup>b</sup> | \( \Delta G_{cis-trans} \) | 19.3 ± 10.8       | [16]|
| Theory         | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) | 1.2 ± 0.1         | [14]|
|                | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) (clockwise)<sup>h</sup> | 18.5 ± 11.8       | [14]|
|                | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) (counterclockwise)<sup>h</sup> | 18.9              | [14]|
|                | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) | 1.2               | [15]|
|                | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) (clockwise)<sup>h</sup> | 12 ± 7            | [15]|
|                | /His-Ala-Gly-Pro-Ile-Ala<sup>b</sup> | \( \Delta F_{cis-tran} \) (counterclockwise)<sup>h</sup> | 12 ± 7            | [15]|

<sup>a</sup>For \textit{in vitro} experiments, \( \Delta F \) is about equal to \( \Delta G \) [10].

<sup>b</sup>Thermodynamic properties for a peptide having the CypA target sequence have not been determined experimentally neither in CypA nor in water solution.

<sup>c</sup>Free energy differences between trans and cis minima at standard condition.

<sup>d</sup>Activation free energies. The experimental \( \Delta G \) values are converted from kinetic data for SUC-Ala-Phe-Pro-Phe-DNa in CypA using Transition State Theory with a prefactor of \( k_B T/\hbar \) (T = 283 K for [17], T = 273 K for [16]).

<sup>e</sup>Free energy differences between trans and cis minima.

<sup>f</sup>SCC-DFTB/TIP3P umbrella sampling calculation. The free energy is calculated as a function of \( \tau(C_1-\delta(C(i)-C(i-1)) \) similar to \( \tau(C_2-\delta(C(i)-C(i-1))) \), see Figure 2.

<sup>g</sup>Amber force field umbrella sampling calculation using \( \omega \) in NVE ensemble (Figure 2).

<sup>h</sup>C\( \alpha \)\( i \) respective to C(i) is in the same side (clockwise) or in the opposite site (counterclockwise) of proline ring (Figure 2).

doi:10.1371/journal.pcbi.1000309.t001

doi:10.1371/journal.pcbi.1000309.g002

**Figure 2. Dihedral angles in prolyl isomerisation.** The dihedral angles \( \omega \) and \( \zeta \) identify the trans \((-\pm 180^\circ)\) or cis \((-0^\circ)\) conformation. The angle \( \psi \) determines the rotation along C\( \alpha(i)-C(i) \) bond. The dihedral angles \( \chi_0, \chi_1, \chi_2, \chi_3 \) and \( \chi_4 \) measures the puckering of the proline ring.

doi:10.1371/journal.pcbi.1000309.g002
the transition state regions \( R \) hereafter). Briefly, we performed a cluster analysis of all the structures [40] and evaluated their population using an umbrella sampling-like approach [41]; minima and TSs are characterized by several clusters and only one cluster, respectively (See Text S1 for details). Free energy barriers were then calculated as differences between the TS cluster and the lowest cluster in the basin of the minimum.

The sampling of TSs was higher in PEPT-WAT than in PEPT-CypA, as shown by the few number of clusters structured in TSs' \( R \) (i.e. less than 10 structures). To improve sampling in PEPT-CypA, we performed additional 4 ns-long MD simulations in which the \( \zeta \) and \( \psi \) angles were harmonically restrained at the values of the TSs. Restraint center and their associated force constants were fitted so as to keep the entire MD within the region of interest.

To test the robustness of our insights on the enzyme mechanism obtained by the above free energy profiles, we compared our results with free energy plots as a function of other pairs of CVs. These are: \( \bar{\theta} \) the \( \zeta \) angle and the proline nitrogen (PAN) pyramidalization \( \bar{\rho} \) that might have a role in peptide prolyl cis/trans rotation [25], \( \rho \) is defined as the distance between P4N and the center of a plane determined by three atoms belonging to G3 and P4 (Figure 3 and Table 2), (ii) The \( \zeta \) and PAN H-bond coordination number \( \langle s \rangle \) that has been shown to be important for in vacuo pyramidalization [25]:

\[
s_1 = \sum_j \frac{1 - (r_{PAN,j} - r_0)^n}{1 - (r_{PAN,j} - r_0)^m},
\]

where \( r_{PAN,j} \) is the distance between PAN and the atom \( j \) (excluding solvent molecules) while \( r_0, n, m \) are parameters chosen to obtain \( s_1 \approx 1 \), when the P4N forms one H-bond (Table S2). \( s_1 \) increases with the number of H-bonds formed by P4N with H-bond acceptors of the peptide (in water) or of the peptide and protein in the complex.

(iii) \( \rho \) and \( s_1 \), to assess the correlation between the pyramidalization and H-bonding of the PAN atom.

Finally, we calculated the free energy profiles as a function of selected pairs specific for PEPT-WAT or for PEPT-CypA. For PEPT-WAT, these are the following: \( \bar{\theta} \) the \( \zeta \) and PAN coordination number with water solvent molecules (\( s_2 \)), \( s_2 \) is defined as \( s_1 \) except that \( j \) runs over the water oxygen atoms. This pair tests whether water is able to catalyze pyramidal isomerization by forming an H-bond with PAN. \( \bar{\theta} \) and \( s_2 \), to evaluate if there is a correlation between PAN pyramidalization and its H-bonding to water molecules.

For PEPT-CypA, we introduced CVs that relate the substrate to the enzyme: \( \bar{\theta} \) hydrophobic coordination numbers \( s_3, s_4 \) and \( s_5 \), i.e. quantities which depend only on non-polar carbon atoms. \( s_3 \), \( s_4 \) and \( s_5 \) describe the hydrophobic interaction between the peptide and the protein:

\[
s_3/4/5 = \sum_{i,j} \frac{1 - (r_{ij} - r_0)^n}{1 - (r_{ij} - r_0)^m},
\]

where \( r_{ij} \) is the distance between atoms \( i \) and \( j \). \( i \) runs over the residues involved in the rotating prolyl bond of the peptide (G3P4) for \( s_3 \), and the C and N terminal of the peptide (H1, A2, I5 and A6) for \( s_4, s_5 \) runs over any non-polar carbon of CypA belonging to residues located at 4 Å or less from the peptide (Table S2) for both \( s_3 \) and \( s_4 \). The parameters \( n \) and \( m \) were chosen so as to distinguish conformations with and without hydrophobic interactions (Table S2). This was done by running 1 ns MD test simulations with different \( n \) and \( m \) parameters.

\( s_5 \) describes a subset of such hydrophobic interactions, which play a particularly important role during the isomerisation, as suggested by NMR studies [21]. These are the interactions of L98 and S99 in CypA and C terminal of the peptide (A5, A6). Thus, for \( s_5 \), \( i \) runs over E5 and A6, while \( j \) runs over the atoms of L98 and S99.

Finally, we introduced the reaction coordinate \( s_6 \):

\[
s_6 = \sum_j \frac{1 - (r_{R55Ns,j} - r_0)^n}{1 - (r_{R55Ns,j} - r_0)^m},
\]

where \( j \) runs over all the H-bond donors/acceptors of the peptide or of CypA that were found within 3 Å from R55 N atoms in 1 ns MD: those turned out to belong to residues Q63 and N149; \( r_{R55Ns,j} \) is the distance from any N atom of R55 and any atom \( j \), \( r_0, n, m \) are chosen as in \( s_1 \) (Table S2). This reaction coordinate increases with the number of H-bonds formed with R55 and hence measures the ability of this residue to form H-bonds with the peptide or CypA residues (organizing its active site): this has been indicated as crucial for the catalysis [14].

The free energy is calculated as a function of \( \bar{\theta} \), \( \zeta \), \( s_3, s_4 \), \( s_5 \); \( \bar{\theta} \), \( \zeta \), \( s_3, s_4 \) and \( s_6 \) to analyze hydrophobic \( (s_3, s_4, s_5) \) and hydrophilic \( (s_6) \) interactions along with the \( \mathrm{cis} \) \( (\zeta \sim 0^\circ) \leftrightarrow \mathrm{trans} \( (\zeta \sim \pm 180^\circ) \) interconversion.

All metadynamics calculations (17 ns for PEPT-WAT and 40 ns for PEPT-CypA) were performed at the physiological temperature of 310 K. The temperature was controlled by a Nose-Hoover thermostat [42], with coupling time constant of 0.05 ps. Electrostatic interactions were assessed using the particle mesh Ewald schemes [37] with 34 wave vectors in each dimension and fourth-order cubic interpolation. Van der Waals interactions were evaluated as specified for the equilibration phase. The time-step was 1 fs. The Gaussians were added with a frequency of 2 GHz and they had a height of 2.5 kJ/mol, the width \( \bar{\theta} \) for each CV is reported in Table S2. The exchange trial frequency was 400 MHz. All bonds were constrained with the LINCS algorithm [43]. We removed the protein centre of mass translation every 10 steps of molecular dynamics. The calculations were performed with a locally modified version of Gromacs 3.3.1 [27,44,45].

### Calculated Properties at Minima and TSs

The number of hydrogen bonds \( N_{HB} \) was calculated in TS and minima \( R \), supposing a Boltzmann distribution within the cluster, as assumed in [46] for enthalpy calculations:

\[
N_{HB}(R) = \sum_{i \in R} \frac{N_{HB}(i)}{N} \exp\left(-\frac{F(i)}{k_BT}\right),
\]

where \( R \) is the region identifying the minima or the TS (See Figure S1), \( N \) is the number of structures belonging to cluster \( i \) (see Figure S1), \( N_{HB}(i) \) is the number of H-bonds and \( F(i) \) is the free energy for cluster \( i \). The H-bond between the acceptor P4N and donors D is assumed to exist if the distance \( \langle D-P4N \rangle \leq 3.2 \) Å and the angle \( \langle P4N-D-H \rangle \leq 70^\circ \).

The populations of puckered conformations were evaluated in terms of \( \chi^2 \) (dihedral angle \( \mathrm{C2\beta-C3\beta-C4\beta-C5\beta} \) in Figure 2); up-puckering is defined when \( \chi^2 > 10^\circ \), planar puckering when \( -10^\circ < \chi^2 < 10^\circ \), down puckering when \( \chi^2 < -10^\circ \) [47].
Mechanism of Action of Cyclophilin A

A  ζ and ψ dihedral angles
B  P4N pyramidalization

C  P4N interaction with water

D  P4N intramolecular H-bond

CVs for PEPT-WAT

G  Coordination of C-term part of PEPT with L98 and S99

E  P4N@PEPT H-bond with CypA residues
H  Hydrophobic interactions of PEPT N- and C-term parts with CypA residues

F  R55 H-bond with PEPT-CypA complex
I  Hydrophobic interactions of G3P4@PEPT with CypA residues

CVs for PEPT-CypA
The population of puckered conformations $P_{x2(c)}$ is defined as:

$$P_{x2(c)} = \frac{\sum_{i=0}^{\infty} \exp \left( -\frac{F(x)}{kT} \right)}{\sum_{i=0}^{\infty} \exp \left( -\frac{F(y)}{kT} \right)},$$  \hspace{1cm} (7)

where $C$ = (up, down or planar puckering) and $i$ is a cluster belonging to this state. The free energy associated with each conformation $C$ is estimated as $F_C = -\frac{1}{kT} \ln P_{x2(c)}$.

The interface coordination number ($IC$) in PEPT-CypA is defined as $IC = \sum_{i=j}^{\infty} \left( 1 - \frac{r_{ij}}{r_{ij}} \right)$, where $i$ is the index of the carbon atoms and $j$ runs over the carbon atoms of CypA hydrophobic residues within 4 Å of G3 and P4 residues.

The internal energy/entropy contributions are evaluated for the restrained MD simulations of each minima and TS. The internal energy contribution to the free energy is calculated as the potential energy averaged over all conformations of the cluster [46], while the entropic contribution is obtained by the standard thermodynamic relation: $\Delta F = \Delta U - \Delta S$.

Principal component analysis (PCA) [48,49] was used to identify large-scale collective fluctuations in minima and TSs restrained dynamics. This analysis was performed on C2C of PEPT-CypA complex, using Gromacs 3.3.1 [44,45].

Errors

The statistical error of the free energy for minima and TSs are estimated as the largest difference between the $F(\zeta)$ values calculated from different 2D free energy profiles (for more details see text S1). The statistical error on the enthalpy of minima and TSs was estimated similarly. The accuracy of the force-field used (Amber [29]) was established by a comparison with quantum chemical results on a model system. We compared the cis-trans isomerization potential energy of a N-acetyl proline methylamide using the Amber force field [29] and DFT calculations at the B3LYP/6-31G(d) [50-53] level of theory (see text S1 for more details on these calculations). Classical calculations were performed with Gromacs 3.3.1 [44,45]. DFT calculations were performed with Gaussian98 [54].

Results

We calculated the free energy $F$ associated with the isomerization of the G3P4 bond in the /HAGPPI/ peptide as a function of the angles $\zeta$ and $\psi$ (defined in Figure 2), which are fundamental collective variables (CVs) to describe the isomerization process [25].

The peptide is either solvated in water (PEPT-WAT) or in complex with CypA, in aqueous solution (PEPT-CypA). Force-field based metadynamics in its bias-exchange variant, in the canonical ensemble, provided the $F = F(\zeta, \psi)$ profile, from which we defined minima and TS regions and clusters within these regions (See Methods Section).

PEPT-WAT

The free energy profile plot shows that the $\text{trans}_{180}$ conformer, in which $\zeta \sim \pm 180^\circ$ and $\psi \sim 0^\circ$, is the absolute minimum (Figure 4 and Table 2). The second $\text{trans}$ minimum is $\text{trans}_{180}$ ($\zeta \sim \pm 180^\circ$ and $\psi \sim \pm 180^\circ$), whose free energy is higher by 3 kcal/mol (Table 2,

---

Table 2. Minima and TSs population in all the CV pairs containing the dihedral angle $\zeta$.

| System   | CV pairs | $\text{trans}_y$ | $\text{cis}_y$ | $\text{trans}_{180}$ | $\text{cis}_{180}$ | TS1 | TS2 | TS3 | TS4 |
|----------|----------|------------------|----------------|----------------------|-------------------|-----|-----|-----|-----|
| PEPT-WAT | Best fit | 94.0%            | 5.0%           | 63.10 -5%           | 6.710 -9%         | 3.910 -10% | 8.010 -9% | 5.710 -11% |
|          | 1        | 95.6%            | 4.1%           | 0.3%                 | 2.010 -3%         | 5.210 -9% | 4.310 -10% | 7.010 -9% | 2.310 -11% |
|          | 2        | 91.8%            | 5.2%           | 2.7%                 | 0.2%              | 5.410 -9% | 2.410 -10% | 1.010 -9% | 4.510 -11% |
|          | 3        | 94.1%            | 4.0%           | 1.8%                 | 1.010 -1%         | 9.010 -8% | 2.110 -9% | 6.310 -9% | 1.510 -10% |
|          | 4        | 90.6%            | 5.6%           | 3.6%                 | 0.1%              | 7.110 -8% | 1.210 -9% | 5.410 -9% | -              |
| PEPT-CypA| Best fit | 5.2%             | 0.2%           | 18%                  | 76.2%             | 7.110 -10% | 4.310 -9% | 1.710 -9% | 1.510 -10% |
|          | 1        | 2.0%             | 0.6%           | 8.4%                 | 89.1%             | 1.510 -9% | 3.010 -9% | 3.110 -11% | 7.810 -11% |
|          | 2        | 10.2%            | 0.1%           | 25.0%                | 64.6%             | 2.110 -9% | 1.710 -9% | -              | 9.310 -11% |
|          | 3        | 15.0%            | 1.1%           | 17.0%                | 66.9%             | 2.210 -9% | 2.610 -9% | -              | 9.710 -11% |
|          | 4        | 4.3%             | 3.7%           | 40.2%                | 52.7%             | 4.910 -10% | 4.110 -9% | 2.410 -11% | 2.110 -10% |
|          | 5        | 0.8%             | 0.3%           | 16.0%                | 76.4%             | 4.610 -10% | 4.210 -9% | 6.410 -11% | 1.610 -11% |
|          | 6        | 0.7%             | 0.5%           | 16.0%                | 76.3%             | 2.810 -9% | 6.310 -9% | -              | 1.710 -10% |
|          | 7        | 3%               | 3.210 -2%      | 19.6%                | 76.8%             | 2.310 -10% | 2.610 -9% | -              | -              |

---

doi:10.1371/journal.pcbi.1000309.g003
Table 3), possibly because the P4N forms a H-bond to the amide group of the adjacent residue only in trans0 (Table S3, see Figure 2 and Figure 1 for atom labelling).

However, the number of waters around the peptide is larger for trans180 than for trans0, since the former has a more extended structure (Figure 4). This water-peptide interaction may not stabilize enough trans180 against trans0 to overcome the stabilization of the latter by the P4N...I5N intramolecular H-bond.

The free energies of the two cis conformers, (cis0; ζ=0° and ψ=0°) and (cis180; ζ=0° and ψ=180°) are 3 and 5 kcal/mol higher than trans0, respectively. The stabilization of trans relative to cis has been ascribed to steric clashes in the trans conformations (see, e.g. [10]), and is affected by intramolecular interactions. In our case, in trans A6 H-bonds to H1, A2 and G3 with higher persistence than cis (Table S3). This is expected to stabilize further the trans conformation. The reason why cis0 is more stable than cis180 is again the presence of the P4N...I5N H-bond only in the first conformation.

As in the case of trans conformers, there are more waters around the peptide in the cis180 than in the cis0 conformation, but water-peptide interactions do not stabilize significantly the cis180 conformation compared to the intramolecular H-bond interaction that stabilize cis0.

The difference in hydration between trans180 and cis180 is reproduced between trans0 and cis0. A direct comparison of our calculation with classical and QM/MM umbrella sampling studies based on the reaction coordinates θ [13] and τ [14], respectively, is not possible, because these calculations do not distinguish between cis and trans conformations at θ=0° and ψ=180°.

The transX→cisX and cisX→transX (X=0°,180°) isomerization free energy barriers range between 4–18 and 11–15 kcal/mol, respectively (Table 3). These values are similar to those obtained by force-field based umbrella sampling calculations performed on the θ variable (Table 1) [13]. However, early quantum gas phase calculations showed that, at variance with the ζ angle, the θ angle alone does not describe properly the proximity to a saddle point conformation [25].

The water shell does not change along the transX→TS6→cisX pathways (Figure 4). This is consistent with the suggested poor solvent reorganization during isomerisation in peptides of similar size [53,56]. However, the solvent could play an important role for the isomerization of peptides smaller than that considered here.

### Table 3. Free energy (kcal/mol) of PEPT-WAT and PEPT-CypA prolyl cis/trans isomerization.

| System      | Pathway along: | TS1 | TS2 | TS3 | TS4 |
|-------------|----------------|-----|-----|-----|-----|
| PEPT-WAT    | trans→cis      | 14  | 16  | 18  | 17  |
|             | cis→trans      | 11  | 14  | 15  | 15  |
| PEPT-CypA   | trans→cis      | 14  | 12  | 18  | 16  |
|             | cis→trans      | 12  | 13  | 16  | 17  |

doi:10.1371/journal.pcbi.1000309.t003

Figure 4. PEPT-WAT. Free energy (kcal/mol) of the /HAGPIA/ peptide in solution as a function of the dihedral angles ζ and ψ (in degrees, showed in the inset). The plot is divided in transition and minima regions, with representative structures (center of the lower free energy cluster within each minima and TS region) for each region explicitly shown (See Text S1 for details), along with water molecules within a shell of 2 Å (red spheres). doi:10.1371/journal.pcbi.1000309.g004
The isomerisation is enthalpy-driven (Figure S2) similarly to that found experimentally on different proline containing peptides [52,53].

Furthermore, proline puckering populations are fully consistent with statistical distributions across peptidyl prolyl bonds in the Brookhaven Protein Databank [47] (see Text S1 and Figure S3).

The lowest pathway is the cis0→TS1→trans0 pathway with an associated barrier of 11 kcal/mol (Figure 4), also shown by in vacuum DFT potential energy calculations of N-acetylproline methylamide isomerization (see Text S1 and Table S4). This value is of the same order of the experimental values for smaller peptides (4 residues, Table 1). The stabilization of TS1 (as opposed to TS2-TS4) may be caused, at least in part, by the larger persistency of the P4N-I5N H-bond (Table S3 and Table S5). The formation of this H-bond, not only for TS1 but also for TS3, has been previously predicted [25]. P4N instead does not interact with the solvent (See Text S1).

Obviously, the cis0→trans0 pathway is favored over the reverse one, since trans0 is the most stable state.

The isomerizations involving states with y,6,180 are higher in free energy because the P4N-I5N intramolecular H-bond is not present in these conformation (Figure 4).

Again, a comparison with previous free energy calculations in aqueous solution is not possible since previous studies did not discriminate between trans/cis0 and trans/cis180. A similar picture emerges from calculations of the free energy as a function of other CV pairs (Table 2). These include the P4N pyramidalization, the key H-bond between P4N and the H-bond donors of the peptide (Figure 3), and P4N hydration (See Methods, Text S1 and Figure S4).

We conclude that trans0 is the most populated state (Figure 4 and Table 2) and that the fastest kinetic process produces this isomer, starting from the most populated cis isomer (cis0).

The presence of the protein alters dramatically the population of the four minima in the F = F(ζ, ψ) plot (Figure 5). The global minimum, and therefore the most populated state, is not a trans configuration: it is cis180 (representative structure in dataset S4). cis stabilization has been already found by AMBER-based and QM/MM free energy calculations as a function of ζ [15] and τ [14] (τ dihedral angle – defined as Cα(i-1)-O(i-1)-Cα(i)-Cα(i+1) – is similar to – Cα(i+1)-O(i)-Cα(i)-Cα(i+1) – used in this work). Trans0 (representative structure in dataset S1) and trans180 (representative structure in Dataset S2) are 1 kcal/mol higher in free energy than cis180 (representative structure in Dataset S3) is scarcely populated (Figure 5 and Table 2).

Cis180 features highly persistent H-bonds between N102@Cy-pA…G3(O/N)@PEPT (Table 4) as well as hydrophobic interactions between G3P4@PEPT and N102, Q63@CypA, A101@CypA, H126@CypA, F113@CypA, M61@CypA, F60@CypA, L122@CypA (Table 5). Similar hydrophobic interactions stabilize trans180, but less persistently. Moreover, this conformation is stabilized by the R55@CypA…P4O@PEPT hydrophilic interaction. As for cis180, F60@CypA stabilizes also trans0 (Table 5). The other CypA hydrophobic residues that stabilize the conformations at y,6,180 do not interact significantly in trans0 (Table 5). At variance with any other minimum, the active site residues I57@CypA and W121@CypA stabilize solely trans0, i.e. the most stable conformation in water. The importance of W121 in the trans conformation was already reported in free energy calculations [15].

N102 stabilizes cis0 as cis180 forming an H-bond to P4O. Several residues of CypA form hydrophobic interactions with cis0. Most of these residues stabilize also other minima (Q63, A101, F113). However, H54@CypA and A103@CypA stabilize only cis0, the minimum with the highest free energy.

Figure 5. PEPT-CypA. Free energy (kcal/mol) of the /HAGPIA/ bound to CypA as a function of the dihedral angles ζ and ψ (in degrees). The plot is divided in transition and minima regions (see Text S1 for details). Representative structures of each minima and TS conformation (center of the lower free energy cluster within each minima and TS region) in PDB format are given in Datasets S1, S2, S3, S4, S5, S6, S7, S8.

doi:10.1371/journal.pcbi.1000309.g005
The lowest free energy barrier is associated with the \( \text{trans}_{180} \rightarrow \text{cis}_{180} \) (counterclockwise) path (Tables 2 and 3). This is also the only isomerization pathway catalyzed by the enzyme relative to \( \text{PEPT-WAT} \) (Tables 2 and 3). Preferential lowering of the counterclockwise N-terminal rotational free energy barrier is consistent with previous classic MD and free energy calculations [15,57,58].

Moreover, as previously noticed [58], N-terminal residue H1@PEPT is exposed to the solvent, while the C-terminal part is anchored to CypA active site (Table S6). TS2 (representative structure in Dataset S6) is stabilized by strong N102N...G3O@PEPT and N102O@G3P4...G3N H-bonds (see Table 4) and by persistent hydrophobic interactions between the peptide and Q63, A101, H126, F113, F60 on CypA (see Table 5).

The enzyme does not decrease the barrier of the reverse pathway (\( \text{cis}_{180} \rightarrow \text{trans}_{180} \)), as CypA stabilizes similarly \( \text{cis}_{180} \) and TS2 (the latter is a bit more stabilized given the higher persistence of hydrophobic and hydrophilic interactions).

In addition, the H-bonds and hydrophobic contacts stabilize TS3 (representative structure in Dataset S7) and TS4 (representative structure in Dataset S8) along with their connected minima (Table 4, Table 2, and Figure 5). We further notice that the H-bonds in TS3 are the same as in TS2, although less persistent (Table 4).

Table 4. Number of H-bonds of CypA residues with G3P4@PEPT along prolyl cis/trans isomerization.

| CypA...G3P4@PEPT H-bonds | trans\(_{0}\) | cis\(_{0}\) | trans\(_{180}\) | cis\(_{180}\) | TS1 | TS2 | TS3 | TS4 |
|---------------------------|---------|---------|-----------|-----------|-----|-----|-----|-----|
| Total                     | 0.1±0.0 | 0.6±0.0 | 1.6±0.0   | 1.3±0.0   | 0.1±0.0 | 1.9±0.0 | 1.5±0.0 | 1.1±0.0 |
| R55...P4O                 | 0±0     | 0±0     | 0±0       | 0±0       | 0±0 | 0±0 | 0±0 | 1±0 |
| N102N-H...P4O             | 0±0     | 0±0     | 0±0       | 0±0       | 0±0 | 0±0 | 0±0 | 0±0 |
| N102N-H...G3O             | 0±0     | 0±0     | 0±0       | 0±0       | 0±0 | 0±0 | 0±0 | 0±0 |
| N102O...G3N               | 0±0     | 0±0     | 0±0       | 0.7±0.0   | 0±0 | 0.9±0.0 | 0.7±0.0 | 0±0 |
| W121...P4O                | 0±0     | 0±0     | 0±0       | 0±0       | 0±0 | 0±0 | 0±0 | 0±0 |
| R148...P4O                | 0.1±0.0 | 0±0     | 0±0       | 0±0       | 0±0 | 0±0 | 0±0 | 0±0 |

Table 5. Interface coordination number\(^a\) of CypA residues with G3P4@PEPT along prolyl cis/trans isomerization.

| CypA Residues | trans\(_{0}\) | cis\(_{0}\) | trans\(_{180}\) | cis\(_{180}\) | TS1 | TS2 | TS3 | TS4 |
|---------------|---------|---------|-----------|-----------|-----|-----|-----|-----|
| Total         | 14.22±0.02 | 19.89±0.03 | 28.94±0.09 | 39.7±0.3 | 18.7±0.2 | 44.01±0.07 | 40.23±0.04 | 38.96±0.06 |
| Q63           | 0.428±0.008 | 4.25±0.02 | 3.06±0.01 | 2.96±0.04 | 0.31±0.01 | 3.94±0.01 | 3.276±0.008 | 3.30±0.01 |
| A101          | 0.107±0.002 | 3.83±0.02 | 2.73±0.02 | 3.35±0.04 | 0.189±0.008 | 4.94±0.01 | 3.829±0.007 | 4.47±0.01 |
| H126          | 0.313±0.007 | 1.586±0.008 | 5.23±0.02 | 6.02±0.05 | 1.08±0.03 | 6.57±0.02 | 6.76±0.01 | 7.68±0.02 |
| F113          | 0.470±0.008 | 1.97±0.01 | 6.77±0.04 | 11.4±0.1 | 0.66±0.03 | 12.79±0.04 | 11.77±0.02 | 2.95±0.01 |
| M61           | 1.31±0.02 | 1.058±0.09 | 3.06±0.02 | 3.19±0.03 | 0.69±0.01 | 3.58±0.02 | 3.210±0.008 | 10.89±0.03 |
| F60           | 3.84±0.04 | 0.410±0.004 | 1.35±0.03 | 5.49±0.04 | 4.4±0.01 | 4.98±0.02 | 4.17±0.02 | 1.38±0.02 |
| L122          | 0.68±0.01 | 0.61±0.01 | 4.08±0.02 | 4.87±0.03 | 2.66±0.05 | 4.53±0.02 | 4.68±0.01 | 4.71±0.01 |
| I57           | 3.19±0.03 | 0.412±0.006 | 0.883±0.008 | 0.597±0.006 | 1.21±0.03 | 0.577±0.002 | 0.542±0.001 | 0.611±0.002 |
| H54           | 0.146±0.002 | 2.14±0.01 | 0.447±0.002 | 0.275±0.002 | 0.0589±0.0009 | 0.418±0.001 | 0.3447±0.0007 | 0.384±0.001 |
| A103          | 0.111±0.002 | 2.34±0.01 | 0.504±0.005 | 0.463±0.005 | 0.150±0.004 | 0.612±0.002 | 0.631±0.001 | 1.49±0.01 |
| W121          | 3.62±0.05 | 0.385±0.003 | 0.830±0.004 | 1.150±0.007 | 7.3±0.1 | 1.048±0.007 | 0.997±0.003 | 1.091±0.004 |

\(^a\)Interface coordination number (IC) is defined as \( IC = \frac{1}{m} \sum_{i=1}^{N} \frac{1}{j-IC} \), where IC is the index of G3P4@PEPT carbon atoms and j runs over the carbon atoms of CypA hydrophobic residues within 4\(\AA\) of G3P4@PEPT.

doi:10.1371/journal.pcbi.1000309.t004
doi:10.1371/journal.pcbi.1000309.t005
Discussion

Our calculations produced the energetics of minima and TS of PEPT-WAT and PEPT-CypA. A critical discussion of the accuracy of the calculations is necessary to proceed with possible insights on the enzyme working mechanism. The source of errors in a free energy calculations are basically of two types [59].

The first is caused by a non-complete sampling (statistical error): our estimated statistical error turns out to be relatively small, less than 1.5 kcal/mol (see Text S1).

The second is caused by the force field. This is expected to be particularly relevant for the TSs, because P4N passes from almost sp² hybridization (resonance of propyl-peptidyl bond) in minima to sp³ at the TSs (the so-called pyramidalization). Notice that since the Amber99 force-field is not designed for having all bonds fixed [29], the isomerization barriers might have been overestimated. However, we expect this effect to bias all barriers in the same direction, so that our conclusions on free-energy differences should not be affected by this choice.

To estimate the error due to the force field, quantum chemical calculations are compared with force field based ones on model systems. Previous work has shown that DFT-B3LYP with 6–31G(d) basis set provide similar results as more expensive (albeit usually more accurate) MP2 calculations on the N-acetyl-N-N'-methyl proline amide [60]. Therefore, here we limit ourselves to a DFT-B3LYP/6–31G(d) calculation on the N-acetyl proline methylamide system, in vacuum, at 0 K, and compare our results with AMBER calculations.

We find small differences for the minima (of the same magnitude of the statistical error, 1.5 kcal/mol, see Text S1). Instead, the differences are 4 kcal/mol or lower for the barriers. We further notice that in three of the reversible pathways (cis→TS1→trans, cis→TS2→trans, cis→TS3→trans), DFT barriers are higher than force field ones, while QM barriers of cis→TS1→trans are slightly lower than the force-field ones (Table S4).

We conclude that our calculations may provide a relatively accurate estimate of the Boltzmann populations of cis and trans free energy minima. In addition, as the errors associated to the barriers are likely to be similar for all pathways, our calculations can be used for qualitative comparisons of the free energy barriers, i.e. to identify the lowest free energy pathways.

Population of Free Energy Minima

In water, trans is by far the most populated specie (Table 2).

The same result for the potential energy has been obtained with DFT (see Text S1) gas phase calculations on N-acetylproline methylamide.

The protein environment changes dramatically the populations of the proline conformers, stabilizing cis with respect to trans, as previously reported by CyPA-SUC-Ala-Phe-Pro-Phe-pNA complex [17] and CyPAK55A-GA [18]. However, our study provides a quantitative estimate of the populations of conformers (Table 2): in presence of the enzyme, three conformers (trans, trans, and cis) are significantly populated. The most populated one corresponds to cis, that is exactly the most unfavorable state in aqueous environment. This finding has never been reported in literature.

Qualitative Comparison of Free Energy Barriers

CyPA accelerates only one interconversion, namely the one from trans to TS2 to cis (Table 3), as already suggested by previous calculations [15]: our calculated free energy barrier decreases by 4 kcal/mol with respect to the peptide in water.

Indeed, as well as in the other pathways, almost no effect is found on the reverse pathway (cis→TS2→trans, 1 kcal/mol, i.e. within the error of the calculations), again in agreement with Refs. [14,15] (Table 4). The trans→TS2→cis pathway is also associated with the lowest free energy barrier. We conclude that this pathway is the most likely in the enzyme, although we cannot establish with high accuracy the free energy barrier of the enzymatic reaction.

Based on our calculation, we propose the following mechanism (Figure 6):

1. The most abundant conformation of the peptide in water is, as it is well known [12], trans conformation. Specifically, in our system it is trans.

2. CyPA sequesters trans and rapidly interconverts it into trans, because the population of the latter is significantly greater in the protein complex. Notice that the process is much faster than the conversion to cis, as the activation free energy associated with them is one order of magnitude smaller.

3. The enzyme catalyzes the isomerization to the most populated minimum, cis, (trans→TS2→cis pathway). This minimum is strongly stabilized by the same PEPT-CyPA interactions as those observed in the most stabilized transition state (TS2), but less persistently (Table 4). Notice that the reverse interconversion from cis to trans is not likely because: (i) the free energy barrier associated with it is larger (indeed, the trans→TS2→cis pathway is the only pathway catalyzed by the enzyme) and (ii) the population of cis is a few times smaller than cis (Table 2). This contrasts with isomerization in water solution, where the population of the most stable trans conformation (trans) is several orders of magnitude higher than the one of the preponderant cis (cis).

4. The enzyme can now either detach cis or interconvert it to cis. However, only in the latter case H34 does interact with the peptide, consistently with the experimental fact that the H54Q mutation decreases the enzymatic activity: the kcat/Km is reduced to the 15% of its wild-type value [20]. Notice that all the enzymatic activity measures, for mutants of polymorphism, do not provide information on kcat or Km alone but only on their ratio. Therefore one cannot conclude whether a mutation affects either kcat or Km or both.

5. cis is released.

Our mechanism provides a first rationale for mutational data which have not been explained so far. The decreased kcat/Km values (Table S1) of the W121A mutant [20] and the E57V variant [19] cannot be explained in terms of loss of TS stabilization. In our mechanism, these two mutations are likely to affect substrate stabilization in the trans conformation, reducing CyPA ability to capture the substrate. In addition, the dramatic decrease of enzyme efficiency in the R55A mutant (0.10% residual activity [20]), whose effect on TS stabilization is controversial, may be, at least in part, a consequence of the reduced population of trans, the reactant of the CyPA catalyzed pathway (trans→TS2→cis).

Next, we analyze the effect of several mutations and polymorphism, so far ascribed only to TS stabilization [14,15,18], as also found here, that might also have an impact on ground state populations.
Indeed, reduced $k_{\text{cat}}/K_m$ values on enzymes where N102 exchanges to T, H and R residues [19] may be not only due to destabilization of the TS, but also of the $cis_{180}$ conformation. Similarly, some CypA mutations also decrease $k_{\text{cat}}/K_m$: (i) F60A [20] may destabilize TS2, $trans_{0}$ and $cis_{180}$; (ii) F113A and H126Q [20] may affect all the species along the $trans_{180}$R$\rightarrow$TS2$\rightarrow$ $cis_{180}$ pathway (Table S1).

$cis_{0}$, proposed to be the most probable final step in our mechanism based on H54Q mutant, is stabilized exclusively by interactions of H54 and A103 with G3P4.

doi:10.1371/journal.pcbi.1000309.g006

Figure 6. Proposed mechanism of action for CypA. CypA sequesters the most abundant conformation in water, $trans_{0}$, that is rapidly interconverted into the most abundant conformer $trans_{180}$. Then, CypA catalyzes the isomerization of the peptide along TS2, producing the mostly populated minimum, $cis_{180}$. The peptide most probably detaches in the $cis_{180}$ conformation. This readily interconverts to $cis_{0}$ once the peptide is in aqueous solution. CypA residues that form important H-bonds (hbb) and hydrophobic interactions (nnb) with the G3P4 moiety are shown. Residues with almost exclusive relevance to each conformation are highlighted in red.
We conclude that our mechanism is consistent with all the mutational and polymorphism data and provides a structural basis for most of them.

**Supporting Information**

**Dataset S1** Representative structure of trans0 conformation (center of the lower free energy cluster within trans0 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s001 (0.21 MB TXT)

**Dataset S2** Representative structure of trans180 conformation (center of the lower free energy cluster within trans180 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s002 (0.21 MB TXT)

**Dataset S3** Representative structure of cis0 conformation (center of the lower free energy cluster within cis0 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s003 (0.21 MB TXT)

**Dataset S4** Representative structure of cis180 conformation (center of the lower free energy cluster within cis180 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s004 (0.21 MB TXT)

**Dataset S5** Representative structure of TS1 conformation (center of the lower free energy cluster within TS1 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s005 (0.21 MB TXT)

**Dataset S6** Representative structure of TS2 conformation (center of the lower free energy cluster within TS2 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s006 (0.21 MB TXT)

**Dataset S7** Representative structure of TS3 conformation (center of the lower free energy cluster within TS3 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s007 (0.21 MB TXT)

**Dataset S8** Representative structure of TS4 conformation (center of the lower free energy cluster within TS4 region) in PDB format.  
Found at: doi:10.1371/journal.pcbi.1000309.s008 (0.21 MB TXT)

**Figure S1** Left: non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and ψ (in degrees, showed in the inset). The plot is divided in transition and minimum regions. Center: each transition and minimum region contains i clusters. Right: the structures within each cluster i are analyzed measuring χ2 and NHB parameters.  
Found at: doi:10.1371/journal.pcbi.1000309.s009 (1.15 MB TIF)

**Figure S2** Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) and enthalpy (kcal/mol) as a function of the dihedral angle ζ (in degrees) obtained by integrating out the variable p. The entropic contribution is given by the difference between the two curves. The isomerization process is mainly enthalpy-driven.  
Found at: doi:10.1371/journal.pcbi.1000309.s010 (0.15 MB TIF)

**Figure S3** (Top) Up, planar and down proline ring puckering conformations. Puckering is defined based on the value of the χ2 dihedral angle ($\alpha$-C, C=O, C=C, δ): up-puckering for $\chi^2>10^4$, planar puckering for $10^{-2}<\chi^2<10^2$, down-puckering for $\chi^2<10^{-2}$. (Bottom) $\chi^2$ distribution within clusters classified as up, down and planar puckering. $\chi^2$ has a bimodal distribution in all clusters with a largest maximum at $\chi^2=40$+ for up clusters, at $\chi^2=-40^+$ for down clusters and with two even peaks at $-40^-$ and $+40^-$ for planar clusters.  
Found at: doi:10.1371/journal.pcbi.1000309.s011 (0.27 MB TIF)

**Figure S4** (A) Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ (in degrees, showed in the inset) and the pyramidalization p. (B) Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ (in degrees, showed in the inset) and the coordination of P4N with peptide H-bond donors (s1). (C) Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ (in degrees, showed in the inset) and the coordination of P4N with peptide H-bond donors (s2). (D) Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the pyramidalization p and the coordination of P4N with peptide H-bond donors (s1). (E) Non catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the pyramidalization p and the coordination of P4N with peptide H-bond donors (s2).  
Found at: doi:10.1371/journal.pcbi.1000309.s012 (2.08 MB TIF)

**Figure S5** (A) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and the pyramidalization p. (B) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and the coordination of P4N with peptide and enzyme H-bond donors (s1). (C) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and the hydrophobic coordination of G3P4 with CypA (s3). (D) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and the hydrophobic coordination of peptide C- and N-termini (H1, A2, I5, A6) with CypA (s4). (E) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the dihedral angles ζ and the hydrophobic coordination of peptide C-term (I5A6) with L89 and S90 (s5). (F) Enzyme catalyzed /HAGPIA/ peptide cis/trans prolyl isomerization: free energy (kcal/mol) as a function of the pyramidalization p and the coordination of P4N with peptide H-bond donors (s1).  
Found at: doi:10.1371/journal.pcbi.1000309.s013 (1.85 MB TIF)

**Figure S6** Normal modes for fluctuations observed on each minimum and the catalytic TS (TS2) of the enzyme catalyzed /HAGPIA/ isomerization. The modes are calculated from principal component analysis (PCA). Relevant modes (high eigenvalue) that are similar to fluctuations reported previously [15] are displayed.  
Found at: doi:10.1371/journal.pcbi.1000309.s014 (2.01 MB TIF)

**Table S1** Isomerase activity of several wild type cyclophilins and CypA mutants.
Mechanism of Action of Cyclophilin A

References

1. Fischer S, Dunbrack RL, Karplus M Jr (1994) Cis-trans isomerization of the proline dipeptide. J Am Chem Soc 116: 11931–11937.
2. Hodel A, Rice LM, Simonson T, Fox RO, Brunger AT (1995) Proline cis-trans isomerization in Staphylococcal cellulase: multi-substrate free energy perturbation calculations. Protein Sci 4: 636–654.
3. Harrar Y, Bellini C, Faure JD (2001) FKBPs: at the crossroads of folding and catalysis. Nat Cell Biol 3: 235–239.
4. Zheng J, Cheng X, Bai J, Guo R (2005) Phosphorylation-specific prolyl isomerization: is there an underlying theme? Nat Cell Biol 7: 1435–1441.
5. Scarlata S, Carter G (2003) Role of HIV-1 Gag domains in virial assembly. Cell Mol Life Sci 55: 423–436.
6. Wulf G, Finn G, Suizu F, Lu KP (2005) Phosphorylation-specific prolyl isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. Nature 430: 243–252.
7. Wulf G, Finn G, Suizu F, Lu KP (2005) Phosphorylation-specific prolyl isomerization: is there an underlying theme? Nat Cell Biol 7: 1435–1441.
8. Gothel SF, Marahiel MA (1999) Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. Cell Mol Life Sci 55: 423–436.
9. Pastorino L, Sun A, Lu PJ, Zhou XZ, Balastik M, et al. (2006) The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-β production. Nature 440: 528–534.
10. Finklestein AV, Prinsh OB (2002) Lecture 2. In: Protein Physics: A Course of Lectures. Amsterdam: Academic Press. pp 13–22.
11. van Holde KE, Jonson WC, Ho PS (1998) Biological macromolecules. In: Principles of Physical Biochemistry. Upper Saddle River, NJ: Prentice Hall. pp 2–67.
12. Bozco DA, Eisenmesser EZ, Pachapsy S, Sundquist WI, Kern D (2002) Catalysis of cis-trans isomerization in native HIV-1 capsid by human cyclophilin A. Proc Natl Acad Sci U S A 99: 5247–5252.
13. Howard BR, Vajdos FF, Li S, Sundquist WI, Hill CP (2003) Structural insights into the catalytic mechanism of cyclophilin A. Nat Struct Biol 10: 475–481.
14. Li G, Cui Q (2003) What is so special about Arg 55 in the catalysis of cyclophilin A? Insights from hybrid QM/MM simulations. J Am Chem Soc 125: 15028–15038.
15. Agarwal PK, Geist A, Gorin A (2004) Protein dynamics and enzymatic catalysis: investigating the peptidyl-prolyl cis-trans isomerization activity of cyclophilin A. Biochemistry 43: 10056–10061.
16. Koffron JL, Kuzmic P, Kishore V, Colon-Bonilla E, Rich DH (1991) Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. Biochemistry 30: 6127–6134.
17. Kern D, Kern G, Scherer G, Fischer G, Drakenberg T (1995) Kinetic analysis of cyclophilin-catalyzed prolyl cis-trans isomerization by dynamic NMR spectroscopy. Biochemistry 34: 13094–13082.
18. Hur S, Breuer T (2000) The mechanism of cis-trans isomerization of prolyl peptides by cyclophilin. J Am Chem Soc 122: 7303–7313.
19. Fanghänel J, Fischer G (2004) Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. Front Biosci 9: 3453–3478.
20. Zydowsky LD, Ettkorn FA, Chang HY, Ferguson SB, Stolz LA, et al. (1992) Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. Protein Sci 1: 1092–1099.
21. Eisenmesser EZ, Bozco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. Science 295: 1520–1523.
22. Agarwal PK (2004) Cis/trans isomerization in HIV-1 capsid protein catalyzed by cyclophilin A: insights from computational and theoretical studies. Proteins 56: 449–463.
23. Agarwal PK (2005) Role of protein dynamics in reaction rate enhancement by enzymes. J Am Chem Soc 127: 15248–15256.
24. Zhao Y, Ke H (1996) Crystal structure implies that cyclophilin predominantly catalyzes the trans to cis isomerization. Biochemistry 35: 7356–7361.
25. Fischer S, Dunbrack RL, Karplus M Jr (1994) Cis-trans isomerization of the proline dipeptide. J Am Chem Soc 116: 11931–11937.
26. Laio A, Parrinello M (2002) Escaping free-energy minima. Proc Natl Acad Sci U S A 99: 12562–12565.
27. Piana S, Laio A (2007) A bias-exchange approach to protein folding. J Phys Chem B 111: 4553–4559.
28. Florin G, Pastore A, Carloni P, Parrinello M (2006) Using metadynamics to understand the mechanism of calmodulin/target recognition at atomic detail. Biophys J 95: 2768–2777.
29. Wang JM, Ciaplik P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 21: 1049–1074.
30. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 63: 712–725.
31. Bashford D, Karplus M (1990) pK a’s of ionizable groups in proteins - Atomic detail from a continuum electrostatic model. Biochemistry 29: 10219–10225.
32. Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, et al. (2003) H+–e−: a server for estimating pKa’s and adding missing hydrogens to macromolecules. Nucleic Acids Res 33: W368–W371.
33. Kelly BN, Kyere S, Kinde I, Tang C, Howard BR, et al. (2007) Structure of the antiviral assembly inhibitor CAP1 complexes with the HIV-1 CA protein. J Mol Biol 373: 355–366.
34. Maloney MW, Jorgensen WL (2000) A free-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. J Chem Phys 112: 8910–8922.
35. Ferrer SE, Zhang YH, Pastor RW, Brooks BR (1995) Constant-pressure molecular-dynamics simulation: the Langevin piston method. J Chem Phys 103: 4613–4621.
36. Petersen MG, Ferguson DM (1998) Constant temperature simulations using the Langevin equation with velocity Verlet integration. Chem Phys 236: 243–252.
37. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, et al. (1995) A smooth particle mesh Ewald method. J Chem Phys 103: 8577–8593.
38. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26: 1781–1802.
39. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21: 1087–1092.
40. Daura X, Gadde M, Jaun B, Seebach D, van Gunsteren WF, et al. (1999) Peptide folding: when simulation meets experiment. Angew Chem Int Ed Engl 38: 236–240.
41. Marinelli F, Pietrucci F, Piana S, Laio A (2008) A kinetic model of Trp-cage folding from multiple biased molecular dynamics simulations. Submitted.
42. Evans DJ, Holian BL (1985) The Nose-Hoover thermostat. J Chem Phys 83: 4069–4074.
43. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18: 1463–1472.
44. Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model 7: 306–317.
45. Berendsen HJC, Vanderspoel D, VandeVonde R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91: 45–56.
46. Piana S, Laio A, Marinelli F, Van Troys M, Bours D, et al. (2008) Predicting the effect of a point mutation on a protein fold: the villin and actin headpieces and their Pro62Ala mutants. J Mol Biol 375: 460–470.
47. Ho BK, Coutsias EA, Seok C, Dill KA (2005) The flexibility in the proline ring couples to the protein backbone. Protein Sci 14: 1011–1018.
48. Garcia AE (1992) Large-amplitude nonlinear motions in proteins. Phys Rev Lett 68: 2696–2699.
49. Amadei A, Linsen AB, Berendsen HJ (1993) Essential dynamics of proteins. Proteins 17: 412–425.
50. Lee CT, Yang WT, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37: 785–789.
51. Becke AD (1993) Density-functional thermochemistry III. The role of exact exchange. J Chem Phys 98: 5640–5652.
52. Vosko SH, Wilk L, Nusair M (1980) Accurate spin-dependent electron correlation energies for local spin density calculations: a critical analysis. Can J Phys 58: 1280–1281.
53. Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev 136: B864–B871.
54. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. J Phys Chem 98: 11623–11627.
55. Frisch MJ, et al. (1998) GImage 98. Pittsburgh, PA: Gaussian Inc.
56. Stein RL (1993) Mechanism of enzymatic and nonenzymatic prolyl cis-trans isomerization. Adv Protein Chem 44: 1–24.
57. Eberhardt ES, Loh SN, Hinck AP, Raines RT (1992) Solvent effects on the energetics of prolyl peptide bond isomerization. J Am Chem Soc 114: 5437–5439.
58. Trzesniak D, van Gunsteren WF (2006) Catalytic mechanism of cyclophilin A as observed in molecular dynamics simulations: pathway prediction and reconciliation of X-ray crystallographic and NMR solution data. Protein Sci 15: 2544–2551.
59. Mark P, Nilsson L (2007) A molecular dynamics study of cyclophilin A free and in complex with the Ala-Pro dipeptide. Eur Biophys J 36: 213–224.
60. McDowell SE, Špačková N, Spooner J, Walter NG (2007) Molecular dynamics simulations of RNA: an in silico single molecule approach. Biopolymers 85: 169–184.
61. Jhon JS, Kang YK (1999) Imide Cis-Trans isomerization of N-Acetyl-V-methyl proline amide and solvent effects. J Phys Chem A 103: 5436–5439.