INSTITUTO POLITÉCNICO NACIONAL

ESCUELA NACIONAL DE CIENCIAS BIOLÓGICAS
SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

DOCTORADO EN BIOMEDICINA Y BIOTECNOLOGÍA MOLECULAR

Molecular simulations of the dynamics and stability of triosephosphate isomerase from *Plasmodium falciparum*

T E S I S

QUE COMO UNO DE LOS REQUISITOS PARA OBTENER EL GRADO DE DOCTOR EN CIENCIAS EN BIOMEDICINA Y BIOTECNOLOGIA MOLECULAR.

Presenta:

Lennane Michel Espinoza Fonseca

Directores de tesis:

Dr. Carlos Wong Ramírez
Dr. José Guadalupe Trujillo Ferrara

Diciembre 2009

Este trabajo se realizó en el Laboratorio de Enzimología del Departamento de Bioquímica de la Escuela Nacional de Ciencias Biológicas y en el Laboratorio de Bioquímica III de la Escuela Superior de Medicina del Instituto Politécnico Nacional, bajo la dirección de los Dr. Carlos Wong Ramírez y Dr. José Guadalupe Trujillo Ferrara, con el apoyo económico del Consejo Nacional de Ciencia y Tecnología.

To the wonderful women who have been my support and my source of inspiration:

Mami, Abue Maruca, Karencita and Asya

I am grateful to Dr. José Trujillo Ferrara and Dr. Carlos Wong for their mentorship, freedom and encouragement that allow me to successfully complete this work.

I would like to thank Dr. David Thomas, who patiently taught me the art of scientific deconstruction –to humbly discard something previously learnt and start something new and exciting from scratch.

I thank my committee: Dr. Isabel Baeza, Dr. Alfonso Méndez, Dr. Miguel Ibáñez and Dr. Jorge González y Merchand. I appreciate all of the helpful discussions and comments you have provided to me.

Behind every successfully completed academic work there are invisible, yet heavy, administrative issues that make it possible. I appreciate all the hard work done by Emma Reyes, who dealt with all the paperwork that made this work possible.

An academic career cannot be conceived without the magic of friendship. I would like to thank my friends for all the "philosophical" conversations, dinners, beers, wines, crazy moments and heated discussions that molded part of my personality: Jazmín García, Carlos Acevedo, Mamura Nasirova, Nadja Dolata, Katti Horova and Jovana Bazerkovska. I am lucky to have such amazing friends!

Childhood is a beautiful period where one is pushed by curiosity to discover a full new world, internal and external. I thank my siblings, Alex, Chris, Karla and Karencita for giving me the opportunity to grow up together, learning together, playing together, laughing and crying together. I thank you all for letting me be more than your brother, a friend, a companion and a father. You can't imagine how important you are in my life and how much I love you all!

I thank my uncles Miguel and Manuel for being with me during the important moments of my childhood. They were very generous to share with me their knowledge of life and passions –from chess to bike riding. I thank my father for bringing me to this world.

Some people say "behind every great man, there is a great woman". Although I do not yet see myself as a great man, I could indirectly claim that statement as true as there has been a great woman behind me: my mother. Her wonderful work started about 29 years ago when I happened to end up in her womb. She raised me as a critical, analytical, open-minded, free person with all the love I haven't seen anywhere else. She transmitted me her experiences, her knowledge, her Beatles and, most of all, her love and passion for life. You know that there are no words that can describe all my love and gratitude!

Last, but not least, I would like to thank Asya for all the patience, tremendous support, encouragement and love that she has been giving me in the past intense years full obstacles that, together, we managed to overcome. I feel very fortunate to have you in my life and be part of yours. Много те обичам!

Table of contents

List of Figures and Tables

Figures

Tables

Alphabetic list of abbreviations

AMBER, Assisted model building with energy refinement

CHARMM, Chemistry at Harvard molecular mechanics

DHAP, dihydroxyacetone phosphate

EPR, electronic paramagnetic resonance

GAP, glyceraldehyde-3-phosphate

GROMACS, Groningen machine for chemical simulations

IC_{50}, half maximal inhibitory concentration

IDPs, intrinsically disordered proteins

IDRs, intrinsically disordered regions

MD, molecular dynamics

NMR, nuclear magnetic resonance

OPLS, optimized potential for liquid simulations

PCA, principal component analysis

PDB, protein data bank

PfTIM, *Plasmodium falciparum* triosephosphate isomerase

PPIs, protein-protein interactions

RMSD, root mean square deviation

RMSF, root mean square fluctuations

SASA, solvent-accessible surface area

TCL, tool command language

TIM, triosephosphate isomerase

VMD, visual molecular dynamics

Abstract

Plasmodium falciparum triosephosphate isomerase (PfTIM), which catalyzes the interconversion between dihydroxyacetone phosphate and glyceraldehyde-3-phosphate, is known to be functional only as a homodimer. Although many studies have shown that the interface Cys13 plays a major role in the stability of the dimer, a few reports have demonstrated that structurally conserved Tyr74 may be essential for the stability of PfTIM dimer. To understand the role of Tyr74, we have performed molecular dynamics (MD) simulations of monomeric and dimeric PfTIM mutated to glycine and cysteine at position 74. Simulations of the monomers revealed that mutant Tyr74Gly does not produce changes in folding and stability of the isolated monomer. Interestingly, comparison of the flexibility of Tyr74 in the monomer and dimer revealed that this residue possesses an intrinsic restricted mobility, indicating that Tyr74 is an anchor residue required for homodimerization. Tyr74 also appears to play an important role in binding by facilitating the disorder-to-order transitions of loops 1 and 3, which allows Cys13 to form favorable interactions with loop 3 and Lys12 to be locked in a favorable position for catalysis. High-temperature MD simulations of the wild-type and Tyr74Gly PfTIM dimers showed that the aromatic moiety of Tyr74 is necessary to preserve the geometry and native contacts between loops 1 and 3 at the interface of the dimer. Disulfide cross-linking between mutant Tyr74Cys and Cys13 further revealed that Tyr74 stabilizes the geometry of loop 1 (which contains the catalytic residue Lys12) and the interactions between loops 1 and 3 via aromatic-aromatic interactions with residues Phe69, Tyr101, and Phe102. Principal component analysis showed that Tyr74 is also necessary to preserve the collective motions in the dimer that contribute to the catalytic efficiency of PfTIM. We conclude that Tyr74 not only plays a role in the stability of the dimer, but also participates in the dimerization process and collective motions via coupled disorder-to-order transitions of intrinsically disordered regions, necessary for efficiency in the catalytic function of PfTIM.

Resumen

Se sabe que la enzima triosafosfato isomerasa de Plasmodium falciparum (PfTIM), la cual cataliza la conversión de dihidroxiacetona fosfato a gliceraldehido-3-fosfato, es funcionalmente activa únicamente en su forma homodimérica. A pesar de que varios estudios han mostrado que el residuo Cys13 juega un papel importante en la estabilidad del dímero, muy pocos estudios han demostrado que el residuo Tyr74, el cual es estructuralmente conservado, es posiblemente esencial para brindar estabilidad a la forma dimérica de la PfTIM. Con el propósito de entender el papel de la Tyr74, se llevaron a cabo simulaciones de dinámica molecular (MD, por sus siglas en inglés) de la forma monomérica y dimérica de la PfTIM, con mutaciones en la posición 74 por glicina y cisteína. Las simulaciones de los monómeros mostraron que la mutación Tyr74Gly no produce cambios en el plegamiento y estabilidad del monómero aislado. De manera interesante se observó que la flexibilidad de la Tyr74 tanto en el monómero como en el dímero es intrínsicamente limitada, lo cual indica que la Tyr74 actúa como residuo de anclaje necesario para la homodimerización. Este residuo también cuenta con un papel importante en la homodimerización al facilitar las transiciones desorden-orden de las asas 1 y 3, lo cual permite que la Cys13 forme interacciones favorables con el asa 3 y a la Lys12 mantenerse restringida en una posición favorable para la actividad catalítica. Las simulaciones de los dímeros silvestres y mutados (Tyr74Gly) a temperaturas elevadas mostraron que el anillo aromático de la Tyr74 es necesario para preservar la geometría y los contactos nativos entre las asas 1 y 3 en la interfase del dímero. El entrecruzamiento por enlace disulfuro entre el mutante Tyr74Cys y Cys13 mostró que el residuo Tyr74 estabiliza la geometría del asa 1 (la cual contiene el residuo catalítico Lys12) así como las interacciones entre las asas 1 y 3 mediante interacciones aromáticas con los residuos Phe69, Tyr101 y Phe102. El análisis del componente principal demostró que la Tyr74 también es necesaria para conservar los movimientos colectivos en el dímero que contribuyen a la eficiencia catalítica de la PfTIM. En conclusión, la Tyr74 no solamente juega un papel en la estabilidad del dímero, sino que también participa en el proceso de homodimerización y en los movimientos colectivos mediante transiciones acopladas desorden-orden de regiones intrínsecamente desordenadas, las cuales son necesarias para la eficiencia catalítica de la PfTIM.

1. Introduction

1.1. Triosephosphate isomerase in the glycolytic pathway

During the first stage of glycolysis, α-D-glucose is enzymatically phosphorylated by ATP and further cleaved in half to yield two 3-carbon molecules: dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (GAP) [1]. However, only GAP can be directly degraded in the subsequent reaction steps of glycolysis. The enzyme triosephosphate isomerase (TIM, EC 5.3.1.1) plays a key role in the first phase of glycolysis by rapidly and reversibly converting DHAP into GAP (**Figure 1**).

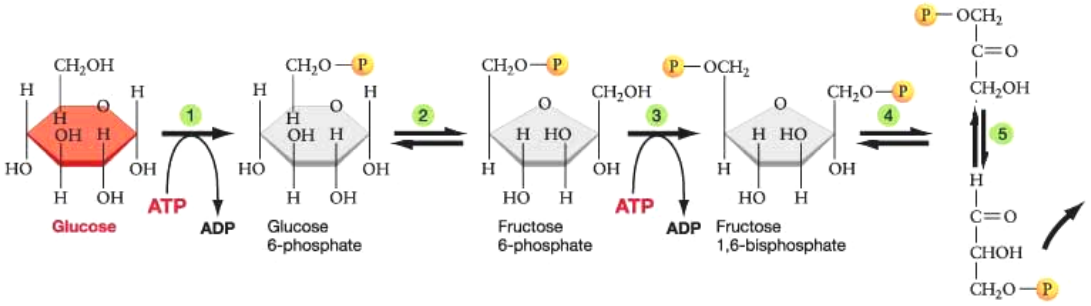

Figure 1. The first stage of glycolysis. The enzymes involved in each step are (1) Hexokinase; (2) Phosphoglucoisomerase; (3) Phosphofructokinase; (4) Fructose bisphosphate aldolase; (5) Triosephosphate isomerase.

TIM has been labeled as a "perfect enzyme" because the chemical steps of the reaction are accelerated sufficiently so that the diffusion steps are rate-limiting [2-4]. Through experimental [5-10] and theoretical studies [11-16], three main paths describing the multistep interconversion of triosephosphates have been proposed (**Figure 2**). The three paths converge to the first step (proton transfer from DHAP to Glu165 and the formation of enediolate) and the last step (proton transfer from the protonated Glu165 to the enediolate to form GAP [17]). Despite these similarities, there are three alternative proposals for the intermediate step (proton from O^1 of the enediolate EDT1 to O^2 in order

to form EDT2, **Figure 2**). Paths A [2-4, 11, 12] and C [7-10, 18, 19] involve generation of an enediol, either by transfer of a proton from and to His 95 (path A) or by transfer of a proton from and to Glu165 (path C). The third mechanism (path B) involves internal proton transfer from O^1 to O^2 without the formation of enediol [13-15].

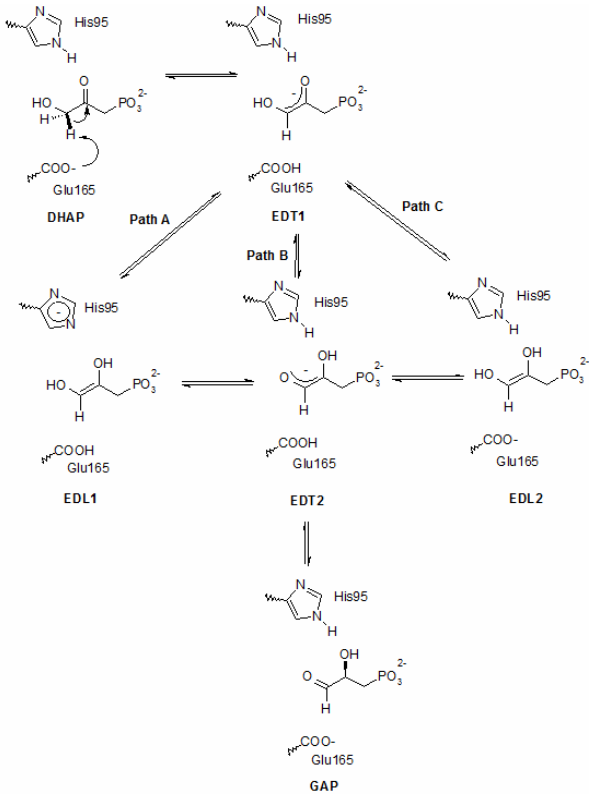

Figure 2. Catalytic reactions for the conversion of DHAP into GAP. EDT1 and EDT2, enediolate 1 and 2, respectively; EDL1 and EDL2, enediol 1 and 2, respectively (Adapted from [17]).

In terms of three-dimensional structure, TIM is one of the best studied enzymes that participate in the glycolytic pathway. For instance, high-resolution x-ray structures of TIM have been solved for *Trypanosoma cruzi* [20], *Trypanosoma brucei* [21], *Leishmania mexicana* [22], *Entamoeba hystolytica* [23], *Plasmodium falciparum* [24], yeast [25], chicken [26], rabbit [27] and human [28]. As this study is focused on the stability and dynamics of *Plasmodium falciparum* TIM (PfTIM), a detailed description of its structure is given bellow.

1.2. The structure of PfTIM

Like most TIMs, PfTIM is functional only as a homodimer. Each monomer, composed of 247 residues, folds as a classical α/β-barrel structure, where eight β-sheets and eight α-helices form the inner and outer layers of the barrel, respectively; these α-helices and β-sheets are connected to each other by loops (**Figure 3**) [24]. The highly conserved active site of PfTIM is located close to the C-terminus of the barrel, as it has been found in all other TIMs. The structure of both subunits is very similar compared to each another, as revealed by superimposition of the Cα trace [24].

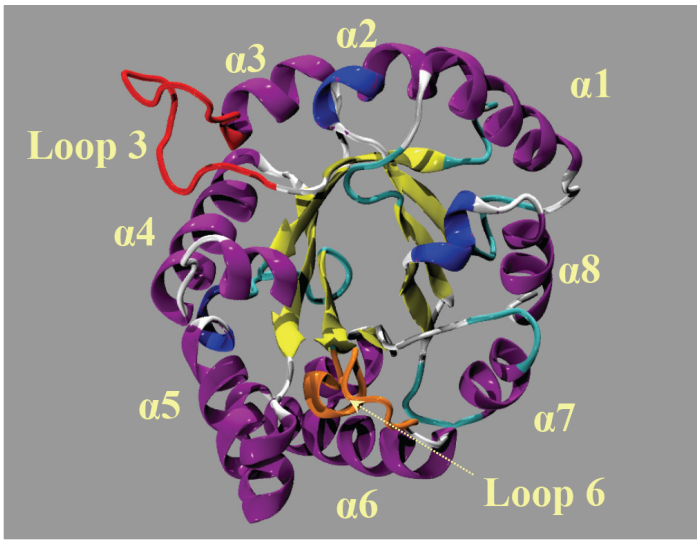

Figure 3. X-ray structure of monomeric PfTIM. α-helices are shown in purple, β-sheets in yellow, turns in cyan and unstructured coils in white. Loops 3 and 6 are shown in red and orange, respectively.

In monomeric PfTIM, β-sheets make an angle of 48–64° with the barrel axis. Consecutive strands of the barrel are at an angle of 12–57°. The helical axes are at angles of 130–144° to the barrel axis; the angle between a β-strand and the associated helix axis varies between 136° and 171°. The core of the barrel is packed with hydrophobic residues of β-sheets (Phe6, Ala8, Val40, Phe42, Gly62, Tyr90, Ile92, Val124, Cys126, Ile161, Val163, Ile207, Gly228 and Leu230). The total van der Waals volume of the residues of

the barrel interior in PfTIM is ~1471 $Å^3$, about 100 $Å^3$ larger than that observed in trypanosomal and human TIM. Moreover, the mean hydrophobicity of PfTIM is larger than that of human and trypanosomal TIMs [24]. The structure of the PfTIM dimer is shown in **Figure 4**.

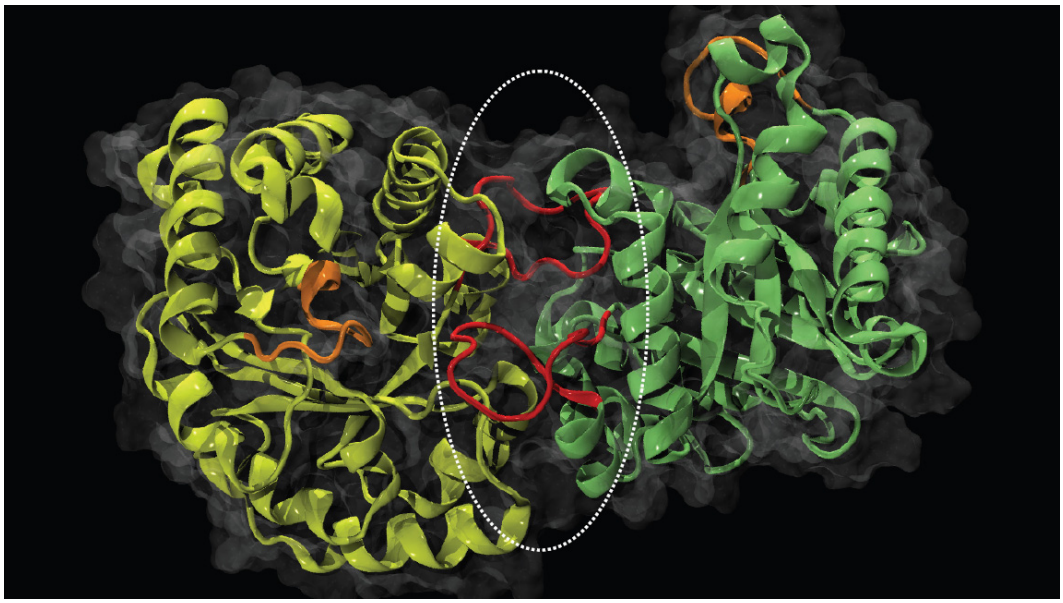

Figure 4. Three-dimensional structure of PfTIM dimer. Individual monomers are shown in yellow and lime. The flexible loop 6 and interface loop 3 are colored in orange and red, respectively. The dashed oval shows the location of the dimer interface.

Dimeric PfTIM possesses two highly conserved active sites, each one located close to the interface of the homodimer. Each active site possesses three catalytic residues: Lys12, His95 and Glu165 (**Figure 5**). Residue Lys12, which provides the positive charge required for substrate binding [29], lies within the less favorable, albeit generously allowed, region of the Ramachandran plot ($\varphi \approx 50^\circ$, $\psi \approx -150^\circ$) [24]. This particular conformation has been hypothesized to play a role in the catalytic efficiency of TIM. The active site possesses a lid that undergoes a large inward movement upon ligand binding [30]; the motions of this lid are also not ligand-gated (i.e., the lid opens and

closes as a natural motion [31]). This lid traps the highly reactive enediol intermediate [32], protects the active site from contact with bulky water [9] and prevents the phosphate elimination reaction [33]. This lid is constituted by loop 6, formed by residues Glu165–Pro178 (**Figure 5**).

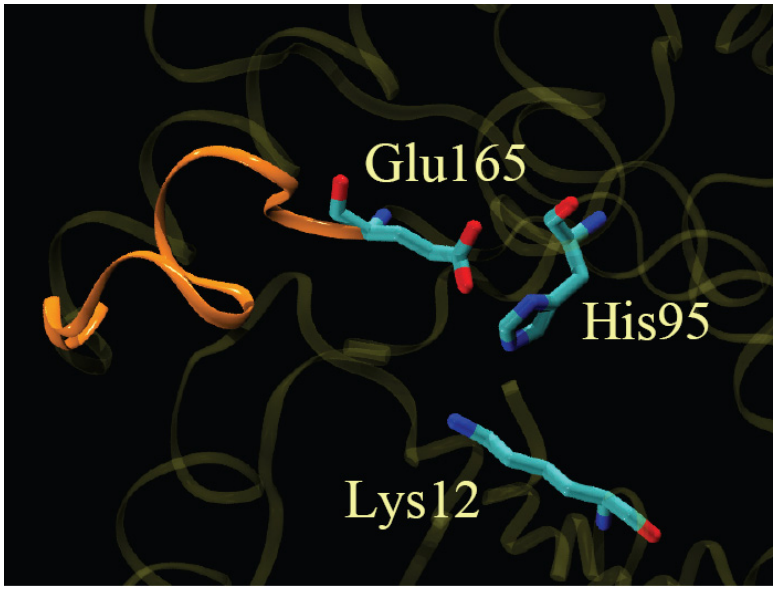

Figure 5. **The active site of PfTIM**. The residues that form the catalytic site are rendered as sticks. The flexible loop 6 is shown in orange.

1.3. Interactions stabilizing the PfTIM dimer

The overall surface area buried in the interface of PfTIM is ~1800 $Å^2$ per subunit, which constitutes 15.5% of the total solvent accessible area of the isolated subunits. In the PfTIM dimer, 12% of the total polar surface and 19% of the total non-polar surface becomes solvent inaccessible upon dimerization [24]. The largest contribution to this buried area comes from loop 3, which interacts with loop 1 (particularly with residue Cys13) of the other subunit (**Figure 6**). Loop 3 (residues 69-79), which protrudes ~13 Å out of the bulk of the monomer and docks to a narrow pocket close to the active site of the other subunit, contributes substantially to intersubunit interactions [24]. This loop is

crucial for dimer stability as ~80% of the intersubunit atom–atom contacts are involved in this region, a feature noted in earlier studies of other TIMs [25, 26, 34].

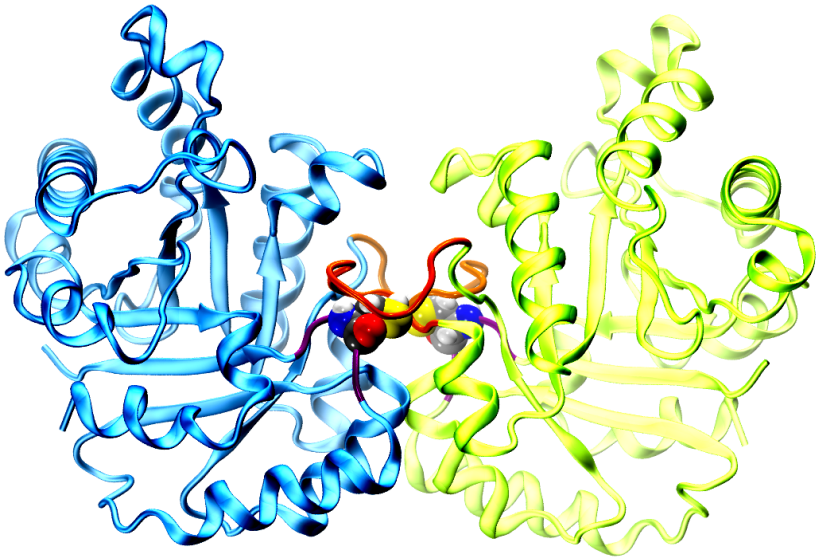

Figure 6. **Principal interactions at the interface of PfTIM**. Loops 1 and 3 are colored as purple and red/orange, respectively; interface Cys13 is shown as van der Waals spheres.

Analysis of the 3-D structure of PfTIM has revealed that the stabilization of the dimer is the result of a combination between polar and non-polar interactions. A large contribution to polar interactions arises from the interaction between Cys13 (loop 1) from one subunit and loop 3 from the other subunit, forming a solvent inaccessible polar cage. In addition to Cys13, this polar cage is constituted by the backbone of Phe69, Asn71, Gly72, Ser73, Tyr74 and Glu77; side chain-side chain interactions are also observed between Cys13 and Ser67, Glu77 and Ser79. Polar interactions around this region have been also observed; for instance, a subunit-subunit salt bridge between Glu77 and Arg98 was found in the x-ray structure of PfTIM [24]. Among the non-polar interactions that stabilize the dimer, a relatively large hydrophobic patch is observed at the interface of

TIM. This hydrophobic patch is formed by residues Leu17, Val44 and Val46 of one subunit and Ile63, Val78, Ile82, Ala83, Leu86 and Ile88 of the adjacent subunit.

Another type of interaction involving the structurally conserved residue Tyr74, aromatic-aromatic one, has also been suggested to participate in the stability of the dimer via the formation of so-called aromatic clusters. However, only a very few studies have addressed the importance of aromatic interactions in the formation, stability and function of PfTIM dimer [35, 36]. A description of the nature of aromatic-aromatic interactions and the contribution of Tyr74-mediated aromatic clusters in the stability of PfTIM is given in the following sections.

1.4. Aromatic-aromatic interactions

π-electrons of an aromatic ring are localized on both sides of the flat ring, inducing a negative quadrupole on the face and a positive quadrupole on the edges. When two aromatic rings are in close proximity to each other, they can bind via quadrupole-quadrupole interactions, giving rise to aromatic-aromatic interactions (also referred to as π-π interactions). In general, there are three main geometrical arrangements of aromatic-aromatic interactions: stacked (or "sandwich"), T-shaped and parallel displaced (**Figure 7**).

Aromatic-aromatic interactions have been studied from both experimental and theoretical standpoints [37]. There is clear evidence that van der Waals and electrostatic interactions play the most important role in the formation, geometrical arrangement and stability of aromatic-aromatic complexes. However, recent studies using the symmetry-adapted perturbation theory (theoretical studies at quantum level) have indicated that

electrostatic, dispersion, induction, and exchange-repulsion contributions are all significant to the overall binding energies, and all but induction are important in determining relative energies [38].

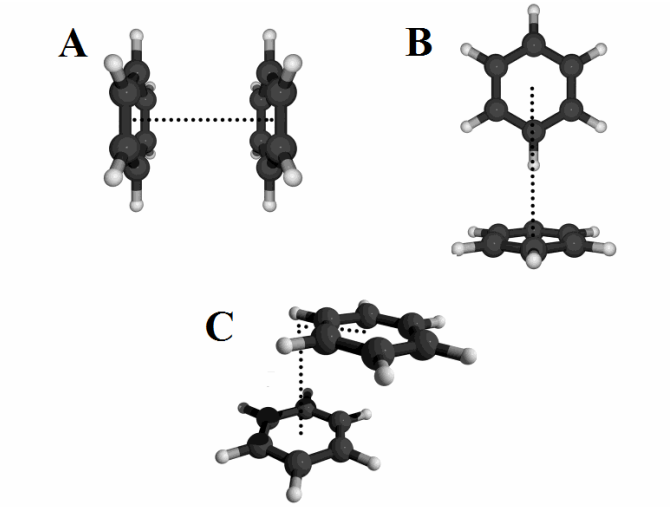

Figure 7. Geometrical arrangement of aromatic interactions between two benzene rings. (**A**) Stacked geometry; (**B**) T-shaped geometry; (**C**) parallel displaced geometry.

Aromatic-aromatic interactions play a key role in many chemical and biological processes. These interactions can influence the stereochemistry of organic reactions [39] and the binding affinities in host-guest chemistry [40]. In biological molecular systems, aromatic residues can engage in specific favorable interactions. For instance, base stacking determines the sequence-dependent structure and properties of DNA as well as recognition of DNA by drugs and regulatory proteins [41, 42]. In other cases, drug-protein recognition heavily relies on aromatic-aromatic interactions; for example, the binding of most of the drugs that reversibly inhibit acetylcholinesterase in the palliative treatment of Alzheimer's disease is driven by this type of interactions [43-48]. Aromatic-aromatic interactions have been shown to be widespread in proteins, providing stability to the secondary, tertiary and quaternary structure of proteins [49-57]. Furthermore, it has

been observed that this type of interactions provide specificity to molecular recognition depending on the environment they take place [58, 59].

1.5. Tyr74-mediated aromatic clusters of PfTIM

In the TIM dimer, there are a total of 40 aromatic residues (20 per monomer). Among those residues only a few aromatic-aromatic pairs are found in the x-ray structure. Particular attention needs to be paid to the so-called "aromatic clusters" located at the interface of the dimer. These clusters have been identified not only in PfTIM, but also in other TIM structures. In PfTIM, these aromatic clusters are formed by Phe69 and Tyr74 from the loop 3 of one subunit and Tyr101 and Phe102 from the adjacent subunit (**Figure 8**).

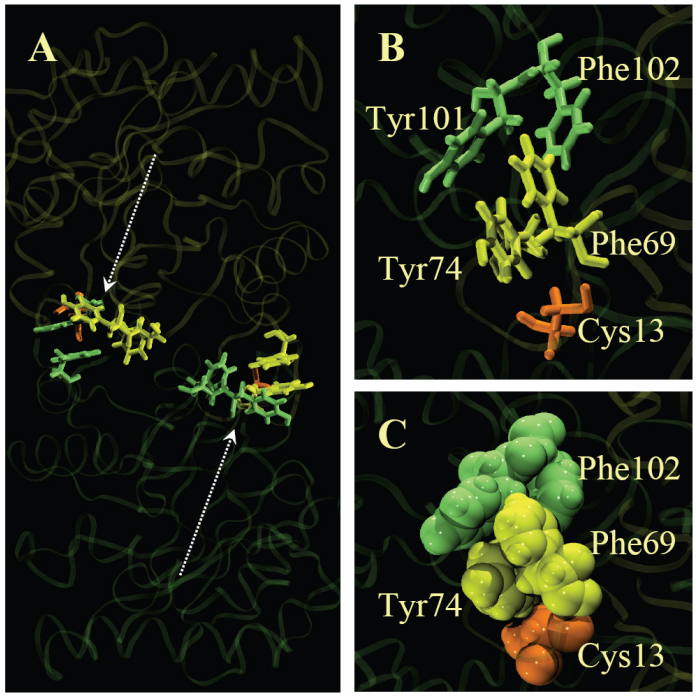

Figure 8. **Structure of the aromatic clusters of PfTIM**. (A) The arrows indicate the position of the aromatic clusters in the PfTIM dimer. (B) Detail of the structure of the aromatic clusters. Residues belonging to the monomer A are shown in lime, while residues from monomer B are shown in yellow. (C) A van der Waals representation of (B), showing the close contacts between the aromatic residues.

Although experimental studies have acknowledged the importance of residue Cys13 [60] as well as loops 1 and 3 [61] for the stability of PfTIM, the role of the structurally conserved interface residue Tyr74 in the stability and function of TIM has not been studied extensively. The backbone of Tyr74 interacts with Cys13 via polar interactions, whereas its side chain participates in two symmetrical aromatic clusters at the interface (**Figure 8A**). These aromatic clusters are formed by Phe69 and Tyr74 from the loop 3 of one subunit and Tyr101 and Phe102 from the adjacent subunit (Figure 8B,C). Gopal et al. [36] studied the mutation of Tyr74Cys with the aim of creating a stable PfTIM dimer via disulfide cross-linking with interface Cys13. Surprisingly, they observed that the oxidized form (cross-linked) of TIM was significantly more stable than its reduced (non-crossed-linked) counterpart, as revealed by thermal precipitation studies. Following up with this study, Maithal et al. [35] engineered the mutation Tyr74Gly in order to determine to what extent the stability of PfTIM dimer is affected by a perturbation of the interface aromatic clusters. The results revealed that the stability and enzymatic activity of the mutant were significantly reduced. Fluorescence and circular dichroism experiments showed that both wild-type and Tyr74Gly PfTIMs displayed similar spectroscopic properties, suggesting that the mutation does not affect folding [35]. Further gel filtration of mutant PfTIM revealed that monomeric and dimeric species are in dynamic equilibrium, with the former predominating at low protein concentration. Low concentrations of urea (< 2 M) were also found to drive the equilibrium of mutant PfTIM toward its monomeric form [35]. More recently, the triple mutant Trp11Phe/Trp168Phe/Tyr74Trp was studied; this mutant possessed a 20-fold reduction in activity compared to the wild type enzyme [62].

The importance of this residue has also been shown in other TIMs. For instance, Gomez-Puyou and co-workers reported a few benzothiazoles, which are able to specifically inhibit trypanosomal TIMs by binding at the dimer interface of the enzyme [63-65]. Further computational studies indicated that the inhibitory mechanism of trypanosomal TIM by benzothiazoles involves the binding to Phe74/Phe75 (equivalent to Tyr74 in PfTIM) and the subsequent destabilization of the aromatic clusters at the interface of the dimer [66-68]. This observation was confirmed by crystallographic studies of the complex benzothiazole-TIM, highlighting the importance of the residue Phe75 in the stability of *T. cruzi* TIM [69].

1.6. Molecular dynamics simulations

Molecular dynamics solves Newton's equation of motion on an atomistic model of a system of *N* particles (i.e., the atoms of a protein) to obtain a trajectory of its motion. Each of these *N* particles will be characterized at time *t* by the corresponding position, velocity and acceleration vectors $\mathbf{r}_i(t)$, $\dot{\mathbf{r}}_i(t)$ and $\ddot{\mathbf{r}}_i(t)$, respectively, where the subscript *i* labels the particle and the usual notation has been used to indicate differentiation with respect to time.

This system is also characterized by a potential energy, $V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$, and a kinetic energy, $T(\dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, \ldots, \dot{\mathbf{r}}_n)$. The potential energy is defined by the potential energy function to be used in the calculations and the kinetic energy is given by:

$$T\left(\dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, ..., \dot{\mathbf{r}}_n\right) = \frac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2(t) \equiv \frac{1}{2}\sum_i m_i \mathbf{v}_i^2(t) \qquad \text{(Eq. 1)}$$

where the summation extends to all the particles in the system, with corresponding masses m_i. In Cartesian coordinates, taking into account that

$$\mathbf{r}_i = x_i\mathbf{i} + y_i\mathbf{j} + z_i\mathbf{k} \qquad \text{(Eq. 2)}$$

$$\dot{\mathbf{r}}_i = \dot{x}_i\mathbf{i} + \dot{y}_i\mathbf{j} + \dot{z}_i\mathbf{k} \qquad \text{(Eq. 3)}$$

$$\dot{\mathbf{r}}_i^2 = \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i = \dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2 \qquad \text{(Eq. 4)}$$

(where **i**, **j** and **k** denote the unit vectors along the x-, y- and z-axes, respectively) the above becomes

$$T\left(\dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, ..., \dot{\mathbf{r}}_n\right) = \frac{1}{2}\sum_i m_i\left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}^2\right) \qquad \text{(Eq. 5)}$$

The Lagrangian function is then defined as:

$$L=T-V \qquad \text{(Eq. 6)}$$

where T is the total kinetic energy, V is the total potential energy and L the Lagrangian. The equations of motion in the Lagrangian form are:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \qquad \text{(Eq. 7)}$$

with one equation for each coordinate q_j (i.e., x_j, y_j or z_j) and associated $\dot{q}_j$; that is, in Cartesian coordinates there are 3*N* equations of motion.

Considering the dependence of the potential and kinetic energy functions, for the *i*-th particle the above equation yields:

$$\frac{d}{dt}\frac{\partial \mathrm{T}}{\partial \dot{\mathrm{x}}_i}-\frac{\partial \mathrm{V}}{\partial \mathrm{x}_i}=\frac{d}{dt}\left(\mathrm{m}_i\dot{\mathrm{x}}_i\right)+\frac{\partial \mathrm{V}}{\partial \mathrm{x}_i}=\mathrm{m}_i\ddot{\mathrm{x}}_i+\frac{\partial \mathrm{V}}{\partial \mathrm{x}_i}=0$$

$$\frac{d}{dt}\frac{\partial \mathrm{T}}{\partial \dot{\mathrm{y}}_i}-\frac{\partial \mathrm{V}}{\partial \mathrm{y}_i}=\frac{d}{dt}\left(\mathrm{m}_i\dot{\mathrm{y}}_i\right)+\frac{\partial \mathrm{V}}{\partial \mathrm{y}_i}=\mathrm{m}_i\ddot{\mathrm{y}}_i+\frac{\partial \mathrm{V}}{\partial \mathrm{y}_i}=0 \qquad \text{(Eq. 8)}$$

$$\frac{d}{dt}\frac{\partial \mathrm{T}}{\partial \dot{\mathrm{z}}_i}-\frac{\partial \mathrm{V}}{\partial \mathrm{z}_i}=\frac{d}{dt}\left(\mathrm{m}_i\dot{\mathrm{z}}_i\right)+\frac{\partial \mathrm{V}}{\partial \mathrm{z}_i}=\mathrm{m}_i\ddot{\mathrm{z}}_i+\frac{\partial \mathrm{V}}{\partial \mathrm{z}_i}=0$$

In vector form, equations (8) may be collected in a single one,

$$\mathrm{m}_i\left(\ddot{\mathrm{x}}_i\mathbf{i}+\ddot{\mathrm{y}}_i\mathbf{j}+\ddot{\mathrm{z}}_i\mathbf{k}\right)=\mathrm{m}_i\ddot{\mathbf{r}}_i=-\left(\frac{\partial \mathrm{V}}{\partial \mathrm{x}_i}\mathbf{i}+\frac{\partial \mathrm{V}}{\partial \mathrm{y}_i}\mathbf{j}+\frac{\partial \mathrm{V}}{\partial \mathrm{z}_i}\mathbf{k}\right) \qquad \text{(Eq. 9)}$$

which can be rewritten as:

$$\mathbf{F}_i\left(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n\right)=\mathrm{m}_i\frac{d^2\mathbf{r}_i}{dt^2}=-\nabla_i \mathrm{V} \qquad \text{(Eq. 10)}$$

Which shows that the force $\mathbf{F}_i$, acting on the *i*-th particle, is derived from the potential V. That is, at any time *t* it must be

$$\frac{d^2\mathbf{r}_i}{dt^2}=-\frac{1}{\mathrm{m}_i}\nabla_i \mathrm{V} \qquad \text{(Eq. 11)}$$

with one such equation for each particle in the system. Thus, the classical behavior of an ensemble of particles is expressed in terms of a set of *N* coupled, second-order differential equations; the designation *coupled* means that the equation for each particle depends on the positions of all the particles in the system.

1.7. Integration of the equations of motion

Solving the equations of motion requires a numerical integration of the differential equations of motion. All the integration algorithms assume the positions **r**, velocities **v** and accelerations **a** can be approximated by a Taylor series expansion:

$$\begin{aligned}
\mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 + \ldots \\
\mathbf{v}(t+\delta t) &= \mathbf{v}(t) + \mathbf{a}(t)\delta t + \frac{1}{2}\mathbf{b}(t)\delta t^2 + \ldots \\
\mathbf{a}(t+\delta t) &= \mathbf{a}(t) + \mathbf{b}(t)\delta t + \frac{1}{2}\mathbf{c}(t)\delta t^2 + \ldots
\end{aligned} \qquad \text{(Eq. 12)}$$

Among the available integration methods, the Leap-Frog algorithm for integration is most popularly used owing to its speed and accuracy. The equations for the Leap-Frog integrator are given as:

$$\begin{aligned}
\mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \mathbf{v}(t+\frac{1}{2}\delta t)\delta t \\
\mathbf{v}(t+\frac{1}{2}\delta t) &= \mathbf{v}(t-\frac{1}{2}\delta t) + \mathbf{a}(t)\delta t
\end{aligned} \qquad \text{(Eq. 13)}$$

and the velocities at time *t* are approximately:

$$\mathbf{v}(t) = \frac{1}{2}\left[\mathbf{v}\left(t+\frac{1}{2}\delta t\right) + \mathbf{v}\left(t-\frac{1}{2}\delta t\right)\right] \qquad \text{(Eq. 14)}$$

In this algorithm, the velocities are first calculated at time $t+\frac{1}{2}\delta t$; these are used to calculate the positions, **r**, at time $t + \delta t$. Thus, the velocities "leap" over the positions, and then the positions "leap" over the velocities.

The smaller the time-step in a simulation, the slower the simulation will run. However, very large time steps can cause instability due to unfavorable conformational jumps. Thus, the integration time step is dictated by the fastest degree of freedom in the

system and must be smaller than the inverse of the fastest frequency of vibration in the system. For polyatomic, bonded systems such as proteins, this is the vibration of 1,2 bonded atoms along the bond axis. A time step of 1 femtosecond is normally used for biological molecules. However, this time step can be increased by using a "constrained dynamics" approach; here, high-frequency vibrations (i.e., those between hydrogen and heavier atoms) are frozen. This approach will be described in the next section.

1.8. Constrained dynamics

It has been shown that constraining the bond lengths between hydrogens and the heavier atoms can speed up the integration of the equations of motion while maintaining the dynamical properties observed in trajectories longer than 1 picosecond. One of the most popular algorithms to achieve such performance improvement is the SHAKE algorithm [70].

When used after each unconstrained time-stepping is performed, the SHAKE algorithm iterates through a list of constraints starting from the first, and adjusts the positions of the atoms as each constraint is being applied. Thus, as the kth constraint, which fixes the bond length between atoms i and j, is about to be applied, any previous k'th constraint with $k' < k$ has already been applied; we call this the previous atom position. If there is no previous constraint, the previous atom position is the position resulting from the unconstrained time-stepping.

The SHAKE algorithm makes the following adjustment to the previous atom position for atom i by:

$$\delta \mathbf{r}_{i,j} = \frac{g\mathbf{r}}{m} \qquad \text{(Eq. 15)}$$

where:

$$g = \frac{d^2 - \mathbf{r}' \cdot \mathbf{r}'}{2\left(\frac{1}{m_i} + \frac{1}{m_j}\right)\mathbf{r} \cdot \mathbf{r}'} \qquad \text{(Eq. 16)}$$

Here, m is the mass of atoms i, j; d is the i–j target bond length for this kth constraint; $\mathbf{r}$ is the distance between atoms i and j before the last unconstrained time step, which satisfies the constraint; $\mathbf{r}'$ be the distance after all k'th constraints with $k' < k$ have been applied.

1.9. Electrostatics with periodic boundary conditions

Periodic boundary conditions allow for a simulation to be performed using a relatively small number of particles in such a way that the particles experience forces as though they were in a bulk solution [71]. This eliminates end effects, and the model is effectively simulated infinitely in x, y and z dimensions. The coordinates of the image particles, those found in the surrounding box are related to those in the primary box by simple translations. Forces on the primary particles are calculated from particles within the same box as well as in the image box. By doing this, it is also ensured that the number of particles N will remain constant throughout the simulation.

When periodic boundary conditions are used in a molecular dynamics simulation, van der Waals and other short-range interactions can be easily calculated within the primary (unit) box. However, electrostatic interactions are long-range and usually reach beyond the neighboring boxes. To calculate the interaction between an ion and its periodic images, the Ewald sum is used [72].

For N charged particles in a cubic unit box of edge length L in a lattice, the electrostatic potential energy among all ions in all copies of the unit box is:

$$V^{zz} = \frac{1}{8\pi\varepsilon_0} \sum_{n}{}' \left(\sum_{i=1}^{N} \sum_{j=1}^{N} z_i z_j \left| r_{ij} + \mathbf{n} \right|^{-1} \right) \quad \text{(Eq. 17)}$$

where z_i and z_j are the charges on the ion, and the sum over **n** is the sum over all lattice points $\mathbf{n} = (n_x L, n_y L, n_z L)$, with n_x, n_y, n_z being integers; the prime excludes summing over the cases with $i = j$ for $\mathbf{n} = 0$. For faster convergence in the calculation of this potential energy, the Ewald summation is used instead:

$$\begin{aligned} V^{zz} = & \underbrace{\frac{1}{8\pi\varepsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\sum_{|\mathbf{n}|=0}^{\infty}{}' z_i z_j \frac{\operatorname{erfc}\left(\kappa \left| \mathbf{r}_{ij} + \mathbf{n} \right|\right)}{\left| \mathbf{r}_{ij} + \mathbf{n} \right|} \right)}_{\mathbf{I}} \\ & + \underbrace{\frac{1}{2\varepsilon_0 L^3} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\sum_{k \neq 0} z_i z_j \frac{1}{k^2} \exp\left(\frac{-k^2}{4\kappa^2} \right) \cos\left(\mathbf{k} \cdot \mathbf{r}_{ij} \right) \right)}_{\mathbf{II}} \\ & + \underbrace{\frac{1}{4\pi\varepsilon_0} \frac{\kappa}{\sqrt{\pi}} \sum_{i=1}^{N} z_i^2}_{\mathbf{III}} \end{aligned} \quad \text{(Eq. 18)}$$

Here, term **I** sums, in real space, the original charges with screening charges that are by construction Gaussian; term **II** cancels the contribution from the screening charges; term **III** corrects the self-energy of the screening charges that is included in term **I**. The value κ is an arbitrary parameter chosen to optimize the summation. In implementation, particle-mesh Ewald [73, 74], an algorithm that scales as $N \log N$, is used.

1.10. Constant temperature and pressure

The native ensemble used in molecular dynamics simulations is the one where the number of particles, the volume and energy remain constant. However, in order to mimic the conditions of biological systems, constant temperature and pressure need to be included in the simulation algorithm. To perform molecular dynamics simulations under constant temperature and constant pressure, the system is coupled to an external bath. The extent of such coupling is determined by a time constant.

During each time step, the velocity of the atoms is proportionally scaled such that the change in temperature over this time-step $\frac{\Delta T}{dt}$ satisfies

$$\frac{dT}{dt} \approx \frac{\Delta T}{dt} = \frac{1}{\tau_T}\left(T_0 - T\right) \quad \text{(Eq. 19)}$$

where T is the temperature of the system, T_0 is the temperature of the bath, and τ_T is the temperature-coupling time constant. The velocity of the atoms v (each with mass m) and the system temperature are related through the kinetic energy through

$$\sum_{\text{atoms}} \frac{1}{2} m v^2 = \frac{3}{2} N k_B T \quad \text{(Eq. 20)}$$

where N is the total number of atoms and k_B the Boltzmann constant. It follows that the velocity of each atom should be scaled from v to λv, where

$$\lambda = \left[1 + \frac{\delta t}{\tau_T}\left(\frac{T_0}{T} - 1\right)\right]^{\frac{1}{2}} \quad \text{(Eq. 21)}$$

Similarly, isotropic pressure coupling is achieved by scaling the boundary sizes and the coordinates of the atoms from l to μl, with

$$\mu = \left[1 - \frac{\delta t}{\tau_p} \left(p_0 - p \right) \right]^{\frac{1}{3}} \quad \text{(Eq. 22)}$$

where p is the system pressure, p_0 the pressure of the external bath, and τ_p the pressure coupling time constant.

1.11. Potential energy function

Theoretical studies of biological molecules permit the study of the relationships between structure, function and dynamics at the atomic level. Considering that most of the problems that one would like to address in biological systems involve thousands of atoms, it is not yet feasible to treat these systems using quantum mechanics. However, the problems become much more tractable when turning to empirical potential energy functions, which are much less computationally demanding than quantum mechanics. In such models, a single nuclear coordinate is used to represent atoms; this approach is justified in terms of the Born-Oppenheimer approximation.

Current generation of potential energy functions (commonly referred to as force fields) provide a reasonably good compromise between accuracy and computational efficiency. They are often calibrated to experimental results and quantum mechanical calculations of small model compounds. Their ability to reproduce physical properties measurable by experiment is tested; these properties include structural data obtained from x-ray crystallography and NMR, dynamic data obtained from spectroscopy and inelastic neutron scattering and thermodynamic data. Among the most commonly used potential energy functions are the AMBER [75], CHARMM [76], and OPLS [77] force fields. The continuing development of force fields remains an intense area of research with

implications for both basic and applied research. We will use the CHARMM (Chemistry at HARvard Molecular Mechanics) force field for our simulations. Over the years, newer versions of the CHARMM force field have been optimized for simulation with biomolecules [78].

1.11.1. The CHARMM potential energy function

The potential energy, V, is a function of the atomic positions **r** of all the atoms in the system; these are usually expressed in term of Cartesian coordinates. The value of the energy is calculated as a sum of bonded terms, E_{bonded}, which describe the bonds, angles and bond rotations in a molecule, and a sum of external or non-bonded terms, $E_{non\text{-}bonded}$, which include the van der Waals and electrostatic terms. The potential energy function can be then written as:

$$\mathrm{V}(\mathbf{r}) = E_{bonded}(\mathbf{r}) + E_{non-bonded}(\mathbf{r}) \qquad \text{(Eq. 23)}$$

1.11.2. Bonded energy terms

The E_{bonded} term is expressed as a sum of three terms:

$$E_{bonded} = E_{bond} + E_{angle} + E_{dihedral} \qquad \text{(Eq. 24)}$$

The first term in the above equation is a harmonic potential representing the interaction between atomic pairs where atoms are separated by one covalent bond, i.e., 1,2-pairs. This is the approximation to the energy of a bond as a function of displacement from the ideal bond length, b_0. The force constant, k_b, determines the strength of the bond. Both ideal bond lengths b_0 and force constants k_b are specific for each pair of

bound atoms, i.e. depend on chemical type of atoms-constituents. The equation representing this term is given as

$$E_{bond} = \sum_{bonds} k_b (b - b_0)^2 \qquad \text{(Eq. 25)}$$

Values of force constant are often evaluated from experimental data such as infrared stretching frequencies or from quantum mechanical calculations. Values of bond length can be inferred from high resolution crystal structures or microwave spectroscopy data.

The second term in above equation is associated with alteration of bond angles θ from ideal values θ_0, which is also represented by a harmonic potential. Values of θ_0 and k_θ depend on chemical type of atoms constituting the angle. These two terms describe the deviation from an ideal geometry; effectively, they are penalty functions and that in a perfectly optimized structure, the sum of them should be close to zero. The mathematical expression for this term is

$$E_{angle} = \sum_{angles} k_\theta (\theta - \theta_0)^2 \qquad \text{(Eq. 26)}$$

The third term represents the torsion angle potential function which models the presence of steric barriers between atoms separated by 3 covalent bonds (1,4 pairs). The motion associated with this term is a rotation, described by a dihedral angle □ and a coefficient of symmetry n around the middle bond. This potential is assumed to be periodic and is often expressed as a cosine function:

$$E_{dihedral} = \sum_{dihedrals} k_\phi (1 - \cos(n\phi)) \qquad \text{(Eq. 27)}$$

In addition to these term, the CHARMM force field has two additional terms; one is the Urey-Bradley term, which is an interaction based on the distance between atoms separated by two bonds (1,3 interaction). The second additional term is the improper dihedral term (see the section on CHARMM) which is used to maintain chirality and planarity of groups.

1.11.3. Non-bonded energy terms

The energy term representing the contribution of non-bonded interactions in the CHARMM potential function has two components, the van der Waals interaction energy and the electrostatic interaction energy. Some other potential functions also include an additional term to account for hydrogen bonds; however, in the CHARMM force field, hydrogen bonds are modeled as a combination of electrostatic and van der Waals interactions. The general expression for the non-bonded energy in the CHARMM force field is:

$$E_{non-bonded} = E_{VdW} + E_{electrostatic} \qquad \text{(Eq. 28)}$$

The van der Waals interaction between two atoms arises from a balance between repulsive and attractive forces. The repulsive force arises at short distances where the electron-electron interaction is strong. The attractive force, also referred to as the dispersion force, arises from fluctuations in the charge distribution in the electron clouds. The fluctuation in the electron distribution on one atom or molecules gives rise to an instantaneous dipole which, in turn, induces a dipole in a second atom or molecule giving rise to an attractive interaction. Each of these two effects is equal to zero at infinite separation r between atoms i and j and become significant as the distance decreases. The

attractive interaction is longer range than the repulsion but as the distance become short, the repulsive interaction becomes dominant. This gives rise to a energy minimum. Positioning of the atoms at the optimal distances stabilizes the system. Both values of energy at the minimum ε_{ij} and the optimal separation of atoms σ_{ij} (which is roughly equal to the sum of van der Waals radii of the atoms) depend on chemical type of these atoms.

The expression for the van der Waals potential is given as:

$$E_{VdW} = \sum_{\substack{non-bonded \\ pairs}} 4\varepsilon_{ij} \left\{ \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right\} \qquad \text{(Eq. 29)}$$

Finally, the electrostatic interaction between a pair of atoms is represented by a Coulomb potential; D is the effective dielectric function for the medium and $\mathbf{r}$ is the distance between two atoms having charges q_i and q_j:

$$E_{electrostatic} = \sum_{\substack{non-bonded \\ pairs}} \frac{q_i q_j}{D\mathbf{r}_{ij}} \qquad \text{(Eq. 30)}$$

1.12. Justification

Despite the biochemical and biophysical data available for the role of Tyr74 in the stability of PfTIM, its exact role in the homodimerization, stability and function of the dimer remains unknown. Thus, complementing existent experimental data with high-resolution molecular dynamics simulations is attractive from two standpoints. First, by analyzing the effect of a perturbation at position Tyr74, it is possible to further understand the role of this residue in the formation, stability and functional dynamics of PfTIM, which in turn will shed new light on how interface Tyr74-mediated aromatic clusters link structure with catalytic function. Second, once we have understood the role of Tyr74-mediated aromatic clusters, this information can be used to rationally design and optimize synthetic molecules that selectively perturb the stability of TIM dimer. By doing so, better drugs with optimized effectiveness and reduced side effects can be developed against parasitic diseases such as malaria and Chagas'.

1.13. Hypothesis

Taking into consideration that Tyr74 is a structurally conserved residue, I hypothesize that Tyr74 possess a crucial role in homodimerization, stability and function of triosephosphate isomerase from *Plasmodium falciparum* via aromatic interactions.

1.14. General and specific aims

1.14.1. General aim

The purpose of this study is to determine whether Tyr74, a key member of the interface aromatic clusters of PfTIM, only plays a role in the stability of the dimer or it partakes in monomer-monomer molecular recognition and functional dynamics required for catalytic activity.

1.14.2. Specific Aims

Aim 1. Evaluate the importance of Tyr74 in the folding and structural stability of monomeric PfTIM.

Aim 2. Understand the role of Tyr74 in the homodimerization mechanism.

Aim 3. Determine the participation of Tyr74 in the structural stability of the interface in the dimer.

Aim 4. Correlate the presence of Tyr74 at the interface with the native collective motions of the dimeric PfTIM.

2. Methods

2.1. Preparation of monomeric PfTIM

The crystal structure of PfTIM (resolution of 2.2 Å) was used as starting coordinates for the simulations (PDB accession code: 1YDV). Considering that the native structure of PfTIM is homodimer in the crystal structure, only the subunit A (as labeled is the crystal structure) was used in this study to simulate the dynamics of monomeric PfTIM. The rationale for choosing subunit A is that there are no structural differences between the two monomers, as revealed by backbone superimposition. The Tyr74Gly mutation was modeled using PSFGEN, a standalone program included in NAMD 2.6 [79]. Side chain ionization states of both monomeric wild-type and Tyr74Gly PfTIMs were adjusted to a pH of 7.0 using the program PROPKA [80]. Hydrogen atoms and $-NH_3^+$ and $-COO^-$ termini were added using PSFGEN. Both monomers were placed in the center of a TIP3P water box with a margin of ~20 Å between the protein and the boundaries of the periodic box. Chlorine and sodium counterions were added to yield a neutral charge on the system, and to produce a physiological ionic strength. Protein, water and ions were modeled with the CHARMM 27 force field [76, 78].

2.2. Preparation of dimeric PfTIM

As we are interested in the native dynamics of the monomer and its perturbation by the mutation of Tyr74 to glycine, we prepared the wild-type and mutated PfTIM dimer. Considering that the native structure of the triosephosphate isomerase is a homodimer, two identical mutations at the interface (i.e., one per monomer) were modeled at the interface of PfTIM. The structure of the wild-type PfTIM was retrieved from the Protein Data Base

(PDB code: 1YDV). The mutation Tyr74Gly was generated with PSFGEN. Side-chain ionization states of the dimers were adjusted to a pH of 7.0 using the program PROPKA [80]. N and C termini were kept as $–NH_3^+$ and $–COO^-$, respectively. Wild-type and mutant dimers were placed in a box of TIP3P water with a margin of ~20 Å between the protein and the boundaries of the periodic box. Chlorine and sodium counterions were added to produce a neutral charge on the system, and to produce a ionic strength of ~150 mM. Protein, water and ions were modeled with the CHARMM 27 force field [76, 78].

2.3. Preparation of disulfide cross-linked and non-cross-linked PfTIM

In accordance with experimental studies [36], we modeled the symmetric mutant Tyr74Cys using the procedure described above. Two individual systems were prepared: one where each Tyr74Cys mutant at the interface is left in its reduced form and one where Cys74 was cross-linked with Cys13 from the adjacent subunit (oxidized form). The ionization states of each dimer were assigned using PROPKA [80], and hydrogen atoms were added with PSFGEN. Each system was embedded in a box of water with margins of ~20 Å. The ionic strength of the systems was further adjusted to ~ 150 mM.

2.4. Molecular dynamics simulations

Molecular dynamics simulations were performed using the program NAMD 2.6 [79]. An NPT ensemble was used for systems studied at 310 K and 333 K, whereas an NVT ensemble was used for systems simulated at a temperature of 400 K; in the latter case, the volume of the systems was adjusted to that obtained at 310 K, and fixed throughout the simulation time. Periodic boundary conditions [71] were imposed on the systems. The

electrostatic term is described by using the particle mesh-Ewald algorithm [73, 74]. The non-bonded cutoff, switching distance and non-bonded pair-list distance were set to 9, 8 and 10.5 Å, respectively. The SHAKE [70] algorithm for bonds to hydrogen atoms allowed a 2 fs time step; the impulse-based Verlet-I multi-step method was used with 2 fs for bonded, 4 fs for short-range non-bonded and 8 fs for long-range electrostatic forces. Constant pressure (for systems simulated under constant pressure conditions) and temperature on the systems are maintained with an isotropic Langevin barostat and a Langevin thermostat, respectively. Thousand steps of conjugate gradient algorithm were used to minimize each system with restraints to protein backbone, followed by 1000 steps without restraints. Systems were warmed up for 60 ps each and equilibrated for 2 ns with lower restraints. Production runs were continued for eight independent systems as described in **Table 1**, for a cumulative total simulation time of 1.3 μs.

Table 1. Summary of the characteristics and conditions of the systems studied.

System	Number of atoms	Number of waters	Temperature (K)	Ensemble used[a]	Length of simulation
Wild-type PfTIM, monomer	75,023	23,681	310	NPT	150 ns
Tyr74Gly PfTIM, monomer	75,024	23,686	310	NPT	150 ns
Wild-type PfTIM, dimer	120,546	37,544	310	NPT	200 ns
Tyr74Gly PfTIM, dimer	120,548	37,548	310	NPT	200 ns
Wild-type PfTIM, dimer	120,546	37,544	400	NVT	150 ns
Tyr74Gly PfTIM, dimer	120,548	37,548	400	NVT	150 ns
Tyr74Cys PfTIM, oxidized	120,186	37,426	333	NPT	150 ns
Tyr74Cys PfTIM, reduced	120,235	37,441	333	NPT	150 ns

[a]NPT, Fixed number of atoms, pressure and temperature; NVT, Fixed number of atoms, volume and temperature

2.5. Analysis of trajectories

2.5.1. Analysis of the secondary structure

In order to track the stability of the secondary structure of the triosephosphate isomerase in the simulations, the STRIDE algorithm to identify the secondary structure motifs will be used [81]. STRIDE is a program to recognize secondary structural elements in proteins from their atomic coordinates. It performs the same task as DSSP by Kabsch and Sander [82] but utilizes both hydrogen bond energy and main-chain dihedral angles rather than hydrogen bonds alone. It relies on database-derived recognition parameters with the crystallographers' secondary structure definitions as a standard-of-truth. The secondary structure analysis is implemented in VMD 1.8.7 [83].

2.5.2. Root mean square deviation

The root mean square deviation (RMSD) is the measure of the average distance between the backbones of superimposed proteins. In the study of globular protein conformations, one customarily measures the similarity in three-dimensional structure by the RMSD of the Cα atomic coordinates after optimal rigid body superposition.

A widely used way to compare the structures of biomolecules or solid bodies is to translate and rotate one structure with respect to the other to minimize the RMSD. Coutsias, et al. [84] presented a simple derivation, based on quaternions, for the optimal solid body transformation (rotation-translation) that minimizes the RMSD between two sets of vectors. They proved that the quaternion method is equivalent to the well-known formula due to Kabsch [85].

The mathematical expression of the RMSD is written as:

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2} \qquad \text{(Eq. 31)}$$

where δ is the distance between N pairs of equivalent atoms (usually Cα). Normally a rigid superposition which minimizes the RMSD is performed, and this minimum is returned. Given two sets of *n* points *v* and *w*, the RMSD is defined as follows:

$$\mathrm{RMSD}(v,w) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(v_{ix}-w_{ix}\right)^2+\left(v_{iy}-w_{iy}\right)^2+\left(v_{iz}-w_{iz}\right)^2} \qquad \text{(Eq. 32)}$$

RMSD for all trajectories was computed with VMD 1.8.7 [83].

2.5.3. Root mean square fluctuations of the Cα trace

The root mean square fluctuations (RMSF) of the Cα atoms from their average positions are of high interest because of their relationship to the atomic temperature factors (B) obtained in x-ray diffraction studies of crystal structures. The mean-square fluctuations can be calculated using:

$$\mathrm{RMSF} = \sqrt{\frac{1}{\mathrm{T}}\sum_{t_j=1}^{\mathrm{T}}\left(x_i\left(t_j\right)-\tilde{x}_i\right)^2} \qquad \text{(Eq. 33)}$$

where T is the time over which one wants to average, and $\tilde{x}_i$ is the reference position of particle *i*. Typically this reference position will be the time-averaged position of the same particle *i*, i.e. $\tilde{x}_i$.

Note that where the difference between RMSD and RMSF is that with the later the average is taken over time, giving a value for each particle *i*. With RMSD the average is taken over the particles, giving time specific values. As for RMSD, RMSF values were computed with VMD 1.8.7 [83].

2.5.4. Fraction of native contacts at the interface

The fraction of native contacts between Cys13 and loop 3 were calculated using TCL scripting in VMD [83]. The fraction of native contacts were calculated using the structure of PfTIM at t=0 ns as a reference. A contact is considered when the distance between Cα atoms of two residues of different chains is ≤ 7 Å.

2.5.5. Ramachandran plots and solvent accessibility

Ramachandran values were extracted using TCL scripting in VMD [83]. Dihedral angles φ and ψ were evaluated separately for each structure in the trajectories and mapped onto a 2D scatter plot. Solvent-accessible surface area (SASA) was calculated for Cys13 side chain in the trajectories. SASA was calculated using a probe radius of 1.4 Å. We automatically evaluated SASA for each conformation in the trajectories using VMD [83].

2.5.6. Principal component analysis

The motions of a protein may be broken down into their principal components by principal component analysis (PCA), also known as essential dynamics analysis [86-88]. Thus, considering only the Cα atoms, the N-residue trajectory can be considered as a vector function of time t, namely

$$\mathbf{r}(t) = \left[r_{1x}(t), r_{1y}(t), ..., r_{Nz}(t) \right]^{\mathrm{T}} \qquad \text{(Eq. 34)}$$

of size $f = 3N$, containing the Cartesian coordinates at time t for residue 1 in the x-direction, residue 1 in the y-direction, . . . , up to residue N in the z-direction. The ij-entry C_{ij} of the covariance matrix C is the covariance of the positions for two degrees of freedom i and j, namely, $C_{ij} = \left\langle \left(r_i(t) - \left\langle r_i \right\rangle_t \right) \left(r_j(t) - \left\langle r_j \right\rangle_t \right) \right\rangle_t$ where $\langle \square \rangle_T$ is the time average over the whole trajectory. Principal component analysis diagonalizes C by solving $\Lambda = \mathrm{T}^{\mathrm{T}}\mathrm{CT}$, so as to obtain the diagonal matrix Λ with the diagonal entries being the eigenvalues ranked by magnitude. The c^{th} column of the transformation matrix T is the c^{th} eigenvector $\mathbf{v}_c$, that is, $\mathrm{T} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_f]$. Principal component analysis (PCA) of the trajectories was performed using GROMACS [89], and generation of the plots describing collective motions was done using the Dinatraj server [90].

2.5.7. Visualization of trajectories and structure rendering

Visualization of the trajectories generated by NAMD was performed using VMD 1.8.7 [83]. Rendering of structures was done with Tachyon, which is included in VMD [83].

3. Results

3.1. Structural stability of monomeric PfTIM

An important feature we address in this study is whether Tyr74Gly mutation can induce considerable perturbation in the native secondary structure content of monomeric PfTIM. To determine the effect of the mutant on the secondary structure, we calculated the fraction of secondary structure averaged over the 150 ns of simulation (**Figure 9**). We found that the structural fractions are almost identical for both wild-type and mutant monomeric PfTIM, clearly indicating that mutation of Tyr74 to glycine does not produce any significant changes in the stability of the secondary structure of the monomer, in agreement with circular dichroism spectroscopy [35].

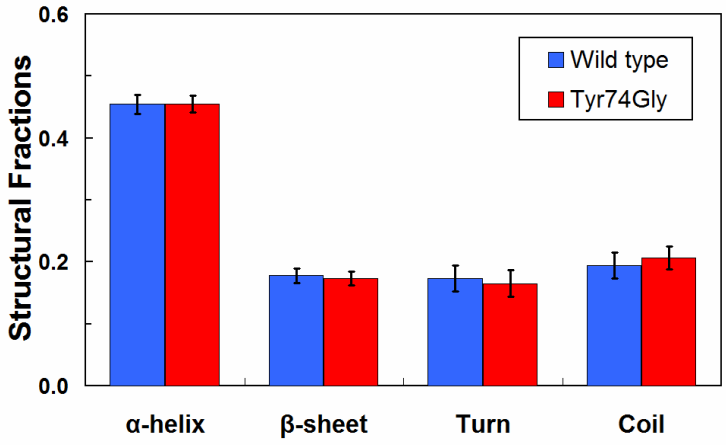

Figure 9. Percentage of secondary structure content for wild-type and mutant monomers.

In order to evaluate the effect of Tyr74Gly on the structural stability and global dynamics of the monomer, we computed root mean square deviation (RMSD) of each structure in the trajectories of the wild-type and mutant PfTIMs. RMSD was calculated by superimposing the Cα trace of each snapshot using the structure at t=0 ns as a

reference (**Figure 10**). RMSD plots showed that both monomers have not extensively deviated from the reference structure, implying that the tertiary structure is largely conserved despite the mutation of Tyr74 to glycine. In addition, RMSD values do not differ much between monomeric wild-type and mutant PfTIM in the 150-ns trajectories, suggesting that in this timescale the mutant Tyr74Gly does not affect the global dynamics of the monomer.

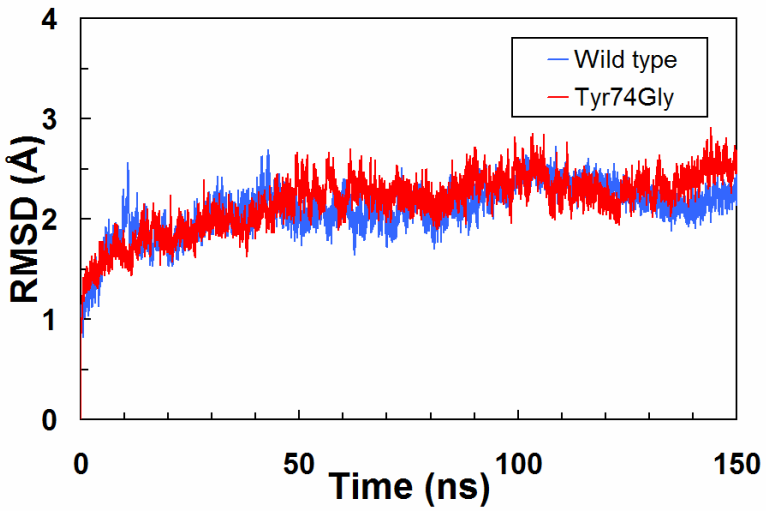

Figure 10. Root mean square deviations of wild-type and Tyr74Gly PfTIM monomers.

Root mean square fluctuations (RMSF) of Cα atoms about their average positions were computed for each residue of wild-type and mutant PfTIM monomers (**Figure 11**). The most mobile parts of wild-type monomeric PfTIM are the connecting loops and the N- and C-termini of the protein. Loops 1, 3 and 6 are especially more flexible compared to the rest of the protein; such increased flexibility is not surprising, as loop 1 and 3 undergo disorder-to-order transitions upon dimerization, whereas the intrinsic flexibility of loop 6 is essential for catalysis [9, 30]. Although replacement of Tyr74 by glycine did not produce any significant change in the secondary and tertiary structure of monomeric

PfTIM, the mutation might produce noticeable changes in the local dynamics of loop 3 (**Figure 11**). Interestingly, Tyr74Gly did not induce important changes in the local flexibility, as the average values of RMSF of loop 3 are 2.45 Å and 2.25 Å for wild-type and mutant monomers, respectively. Likewise, Tyr74Gly did not seem to appreciably affect the mobility of Cα atoms of other regions compared to the wild-type monomer.

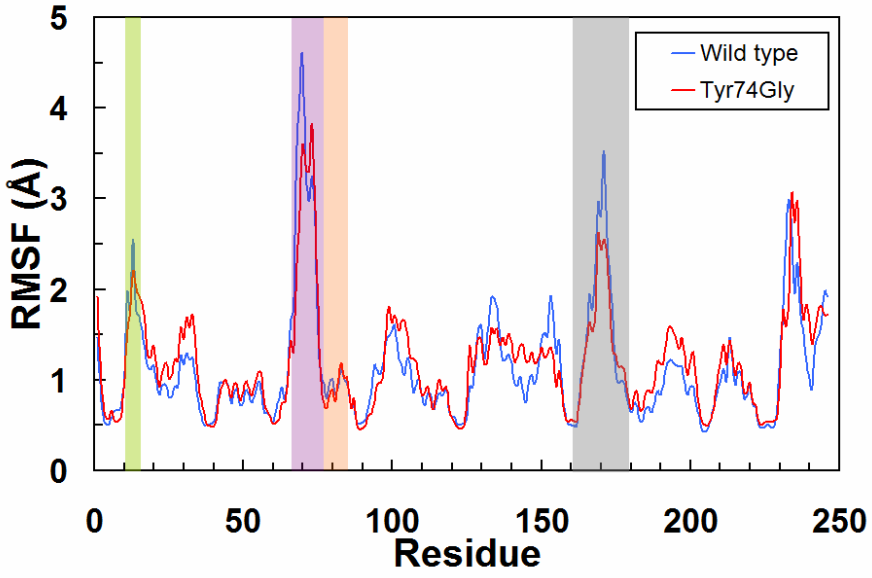

Figure 11. **Root mean square fluctuations of Cα atoms about their average positions computed for wild-type and Tyr74Gly PfTIM monomers**. Green, purple and grey shading indicate the location of loops 1, 3 and 6, respectively.

3.2. Effect of the mutation Tyr74Gly on the structure of the dimer

Experiments have shown that Tyr74Gly significantly reduces the stability of the dimer [35], although the exact destabilization mechanism is not yet clear. Therefore, we performed MD simulations of wild-type and Tyr74Gly dimers. Analysis of the fraction of secondary structure showed that mutation of Tyr74 to glycine does not significant affect the folding of the subunits of dimeric PfTIM (Table 2). Likewise, the tertiary structure of each subunit remained unchanged during the 200 ns of simulation. Time-dependent

RMSD was calculated for wild-type and mutant dimers by superimposing the Cα trace of each snapshot using the coordinates of equilibrated dimers as a reference (**Figure 12**). RMSD of the wild-type dimer did not largely deviate from the original structure (<RMSD> = 2 Å) in this timescale. Similarly, the quaternary structure of the mutated dimer remained fairly stable between 0 and 140 ns (<RMSD> = 2 Å); however, a sudden increase of 1 Å in the RMSD was observed in the mutant at *t*=145 ns.

Table 2. Percentage of secondary structure content for wild-type and mutant dimers at 310K.

	Wild-type (Monomer A)	**Wild-type (Monomer B)**	**Tyr74Gly (Monomer A)**	**Tyr74Gly (Monomer B)**
α-helix	45 ± 2	46 ± 2	45 ± 2	45 ± 2
β-sheet	18 ± 1	17 ± 1	17 ± 1	17 ± 1
Turn	18 ± 3	17 ± 2	18 ± 3	17 ± 2
Coil	19 ± 2	20 ± 2	20 ± 2	21 ± 2

To determine if this increase in the RMSD is the result of a conformational perturbation at the interface due to the removal of Tyr74, we calculated the RMSF of each residue and subunit of wild-type and mutant dimers (**Figure 13**). Compared to the monomeric form of PfTIM (**Figure 11**), loops 1 and 3 of each subunit of the wild-type dimer are considerably less flexible, although active site loop 6 conserves its flexible nature. These structural properties inherent to the wild-type form of the dimer are not observed when the mutation Tyr74Gly is introduced at the interface. This mutation increases the flexibility of loop 1 in both subunits compared to the wild-type form. In fact, the structural fluctuations of loop 1 induced by this mutation are comparable to those observed in the isolated monomers. Likewise, Tyr74Gly induces a shift in the

backbone fluctuations of loop 3 of subunit B (**Figure 13B**). Tyr74Gly also induces an appreciable increase in the dynamics of loop 6 of subunit B.

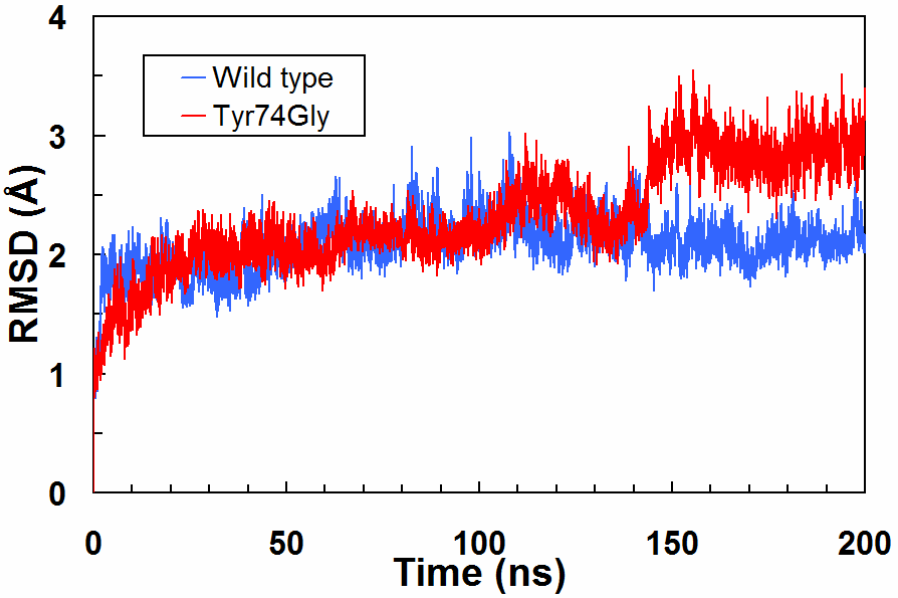

Figure 12. Root mean square deviations of wild-type and Tyr74Gly PfTIM dimers simulated at T=310K.

3.3. Perturbation of native contacts at the interface induced by Tyr74Gly

To monitor the effect of Tyr74Gly on intermonomeric contacts between loops 1 and 3, we calculated the fraction of intermolecular native contacts between Cys13 (loop 1) and loop 3 of adjacent subunit (Q_{Cys13}, **Figure 14**). We chose Cys13 as (a) in the native state of the dimer, it makes critical contacts with loop 3 and (b) its side chain becomes entirely buried upon dimerization.

At physiological temperature, contacts between Cys13 and loop 3 in the wild-type dimer are held together by favorable electrostatic and van der Waals interactions ($\langle E_{int} \rangle \approx$-20 kcal/mol). Given the inherent flexibility of loop 3, the value of Q_{Cys13} is expected to fluctuate in the trajectory of the wild-type dimer; **Figure 14A and B** (blue line) show continuous shifts in time-dependent Q_{Cys13}, in agreement with this assumption.

Nevertheless, residue Cys13 of subunits A and B retains at least 50% of its native contacts with loop 3 (Q_{Cys13}>0.5) during the entire 200 ns of simulation. Conversely, when Tyr74 is replaced with glycine, only 52% (subunit A) and 69% (subunit B) of the total simulation time satisfies a Q_{Cys13}>0.5 (**Figure 14, A and** B, red line).

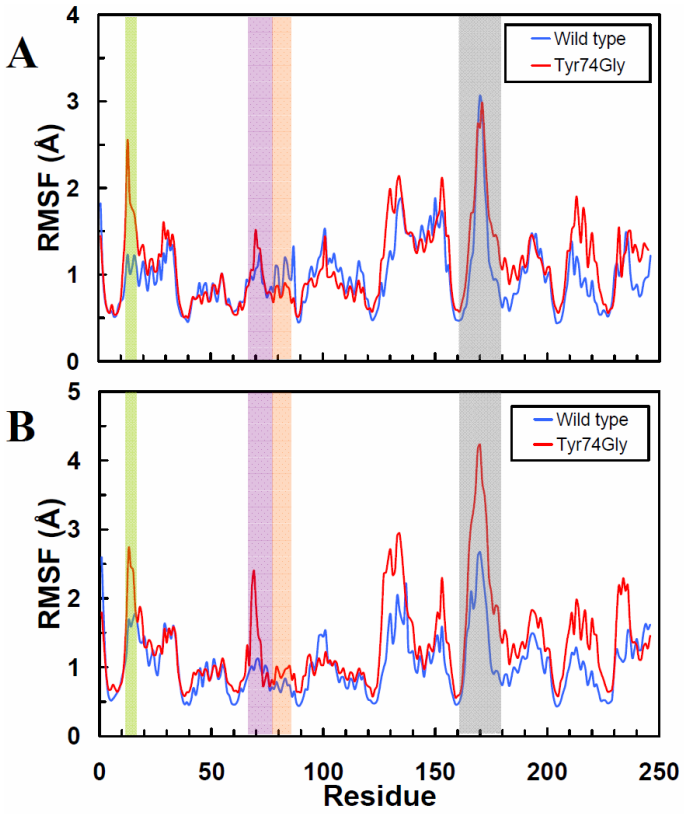

Figure 13. **Root mean square fluctuations of Cα atoms about their average positions computed for (A) subunit A and (B) subunit B of wild-type and Tyr74Gly PfTIM dimers**. Green, purple and grey shading indicate the location of loops 1, 3 and 6, respectively.

High-temperature (400K) 150-ns MD simulations of the dimer were used to accelerate the possible order-to-disorder transitions induced by Tyr74Gly. Despite the use of high temperature, both wild-type and mutant dimers retain most of their secondary and tertiary structure in this timescale. Furthermore, we observed that high temperature did not produce dramatic changes in the RMSF pattern of most of the protein compared to

that obtained at 310K. Nonetheless, the RMSF of loops 3 is 1-1.5 Å larger in the mutant compared to the wild-type, indicating that flexibility of this loop is increased when the side chain of Tyr74 is removed. This increase in flexibility is the result of a large loss of intermolecular contacts between Cys13 and loop 3 (**Figure 15, A and B**). For instance, only 5% (subunit A) and 20% (subunit B) of the total simulation time at 400K satisfies a Q_{Cys13}>0.5, whereas native contacts are largely conserved in the wild-type PfTIM (76% and 97% of the total simulation time for subunits A and B, respectively). These observations also correlate with the increase of solvent-accessible surface area (SASA) of Cys13 in the dimeric mutant (**Figure 16**).

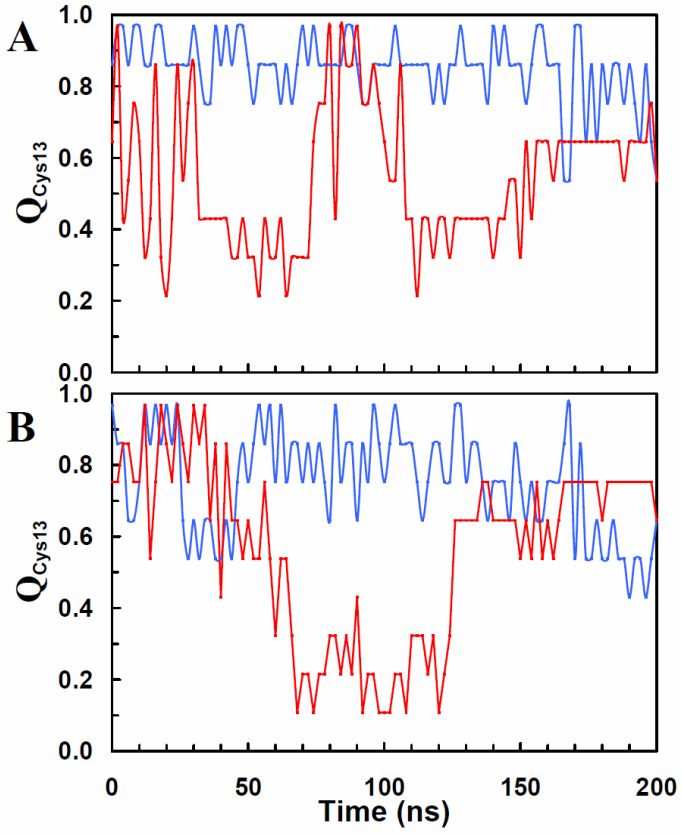

Figure 14. **Fraction of intermolecular native contacts between Cys13 and loop 3 computed for wild-type (blue line) and Tyr74Gly (red line) PfTIM dimers at 310K**. (A) Fraction of intermolecular contacts involving Cys13 of subunit A; (B) Fraction of intermolecular contacts involving Cys13 of subunit B.

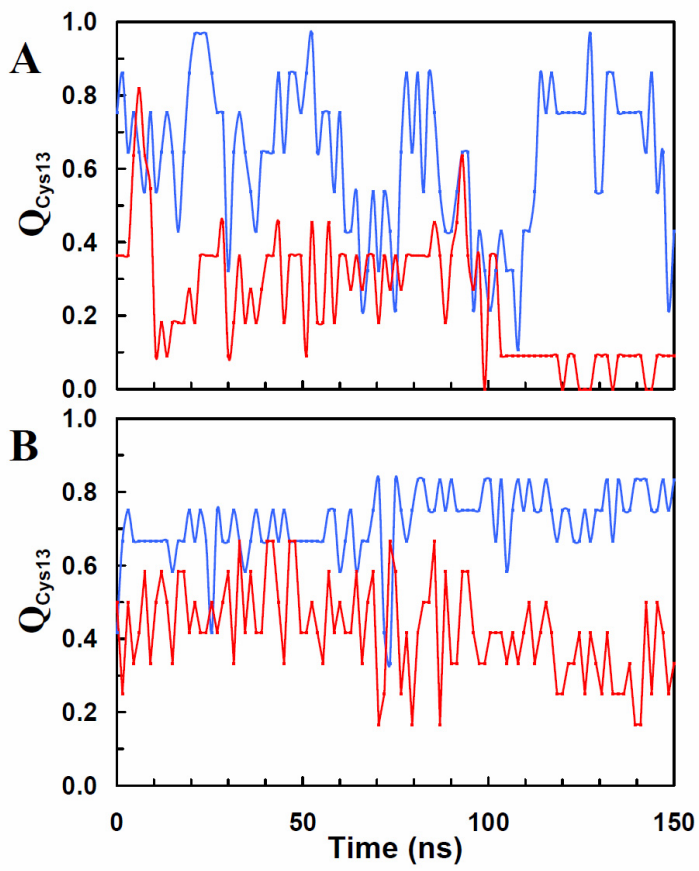

Figure 15. **Fraction of intermolecular native contacts between Cys13 and loop 3 computed for wild-type (blue line) and Tyr74Gly (red line) PfTIM dimers at 400K**. (A) Fraction of intermolecular contacts involving Cys13 of subunit A; (B) Fraction of intermolecular contacts involving Cys13 of subunit B.

3.4. Mutation-induced alteration in the backbone conformation of Lys12

The x-ray structure of PfTIM dimer has revealed that active site residue Lys12 lies within a less favorable, albeit generously allowed, region of the Ramachandran plot ($\varphi \approx 50^{o}$, $\psi \approx -150^{o}$) [24]. As similar Ramachandran values have also been observed in other TIM structures, it is assumed that this particular conformation is required for catalytic function. As the structure of the region surrounding Cys13 is affected by the mutant

Tyr74Gly, we analyzed the mutation-induced alteration in the backbone conformation of Lys12.

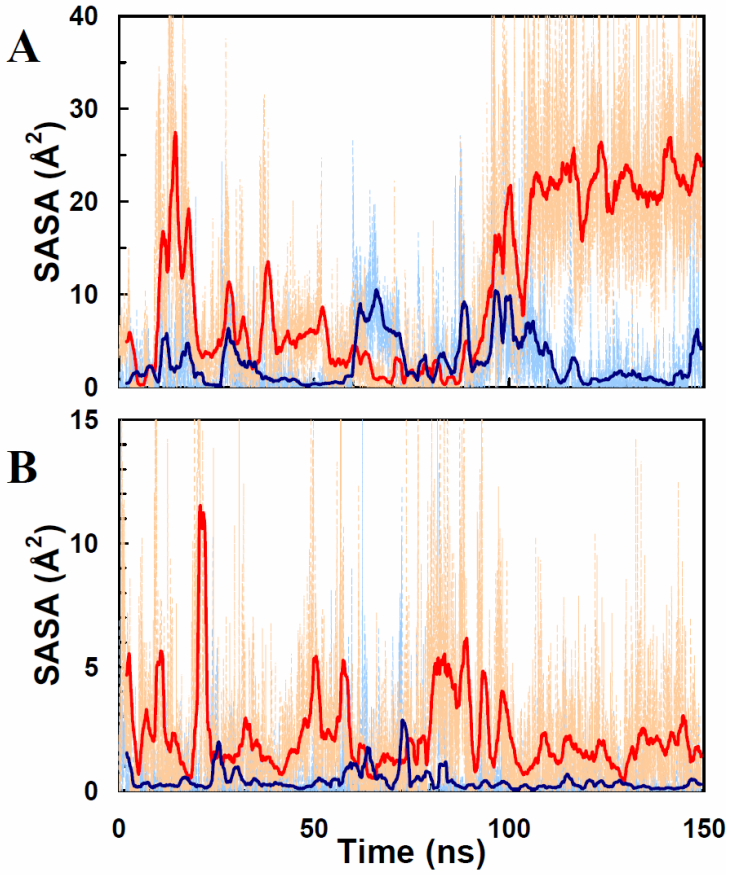

Figure 16. **Solvent-accessible surface area of Cys13 calculated for wild-type (blue line) and Tyr74Gly (red line) PfTIM**. (A) and (B) correspond to Cys13 of subunits A and B, respectively.

First, we mapped φ and ψ angles of Lys12 onto the Ramachandran plot for each structure in the trajectory of wild-type monomeric PfTIM (**Figure 17**). Analysis of the Ramachandran plot revealed the existence of three populated clusters. The first cluster (Figure 6, orange oval) corresponds to the Ramachandran values observed in the crystal structure of the dimer; this conformation remains fairly stable in the trajectory for the first 24 ns. After this period of time, the conformation of Lys12 shifts rapidly to populate a second cluster in the right-handed α-helical region (φ≈-60º, ψ≈-60º) and the so-called

"bridge region" (**Figure 17**, green oval) [91]. Lys12 spends the last 30 ns of simulation populating a backbone conformation having dihedral angles of φ≈-150º and ψ≈0º (Figure 6, purple oval). All these conformational transitions fall within the limits of the energetically allowed regions of the Ramachandran plot. It is worth mentioning that after the initial 24 ns of simulation, Lys12 does not adopt a conformational similar to that in the dimer (φ≈50º, ψ≈-150º). This observation indicates that, in spite of being energetically allowed, this particular backbone conformation does not natively exist in the monomeric form of PfTIM.

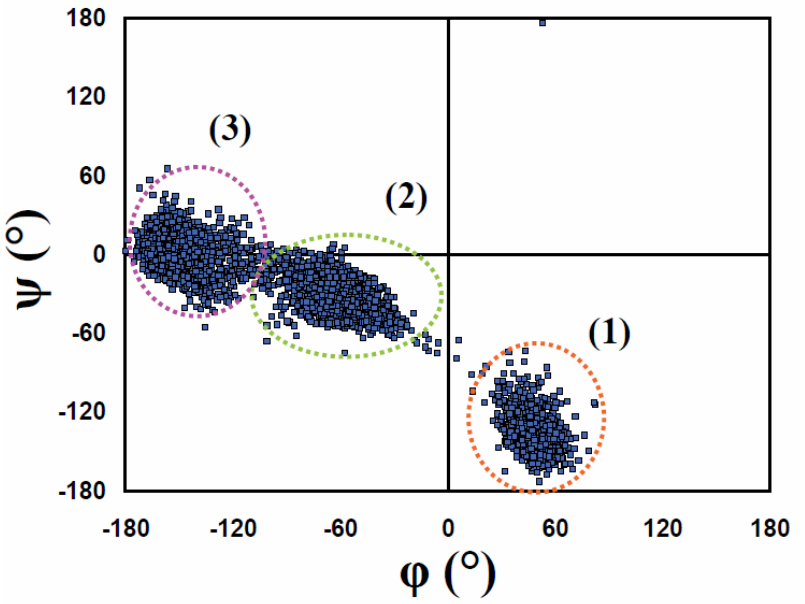

Figure 17. Ramachandran plot of residue Lys12 in the wild-type monomeric PfTIM. Each oval represents a particular population of backbone conformations in the trajectory (see text for interpretation).

Lys12 showed similar conformation patterns in the simulations of wild-type and Tyr74Gly dimers at 310K. The backbone dihedral angles of Lys12 were φ≈50º and ψ≈-150º in each subunit of the wild-type and mutant dimer. Although the free energy of the unbound-to-bound-like conformational transition of Lys12 in the monomer is rather small (ΔG≈1 kcal/mol), local stabilizing interactions at the interface of the dimer may increase

this free energy difference. As this energy barrier may not be overcome at T=310K in the sub-microsecond timescale, structural transitions of Lys12 in the dimer were accelerated using MD simulations at T=400K. Despite the high temperature, Lys12 of each subunit in the wild-type dimer was able to retain its native bound-like conformation (**Figure 18A and B**), in agreement with the results obtained for the wild-type dimer at 310K (φ≈50º, ψ≈-150º). A similar result was obtained for Lys12 of subunit B in the dimeric mutant (**Figure 18D**); however, the removal of Tyr74 side chain induces a change in the backbone conformation of Lys12 similar to that observed in the free monomeric wild-type (**Figure 18C**).

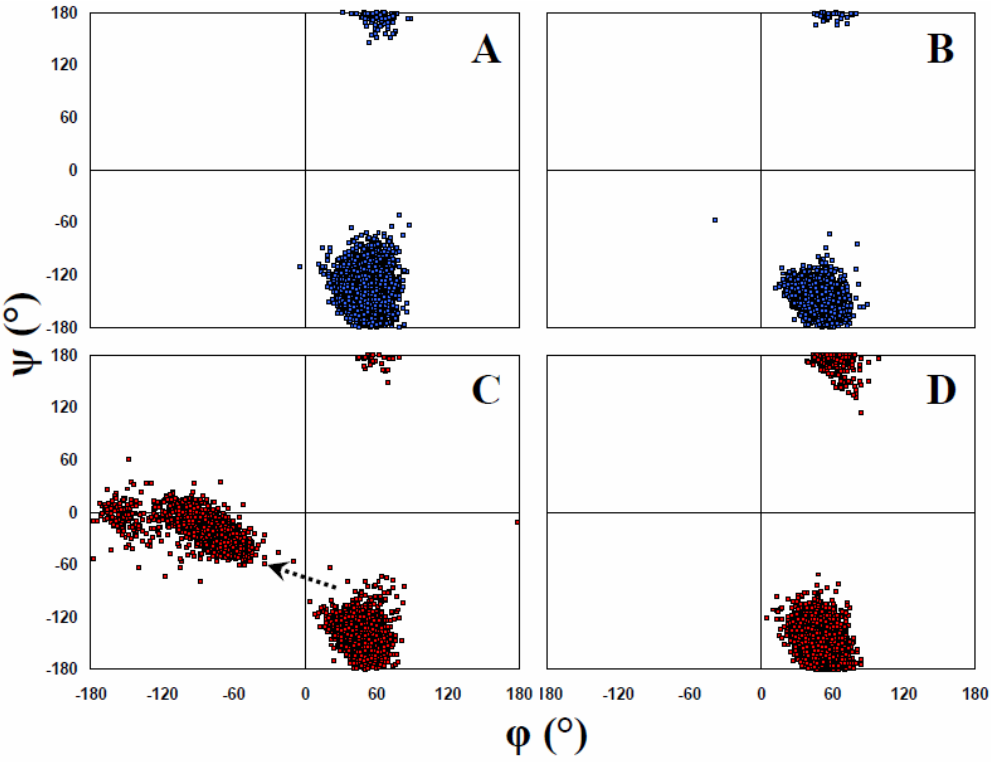

Figure 18. **Ramachandran plot of residue Lys12 in dimeric PfTIM**. (A) and (B) correspond to Lys12 of subunits A and B of wild-type PfTIM dimer, respectively. (C) and (D) correspond to Lys12 of subunits A and B of Tyr74Gly PfTIM dimer. The arrow in (C) indicates the conformational transition of Lys12 in the simulation.

3.5. The effect of disulfide cross-linking at the interface

It has been observed that the mutation of Tyr74 with cysteine followed by chemical cross-linking with Cys13 produces a PfTIM with similar stability to that observed for the wild-type form of the dimer [36]. This observation suggests that Tyr74 somehow participates in the stability of the interface. To address this issue, we performed two MD simulations of dimeric Tyr74Cys mutants: one where Cys13 and Cys74 are connected by a disulfide bond (oxidized PfTIM) and one where Cys13 and Cys74 are not cross-linked (reduced PfTIM).

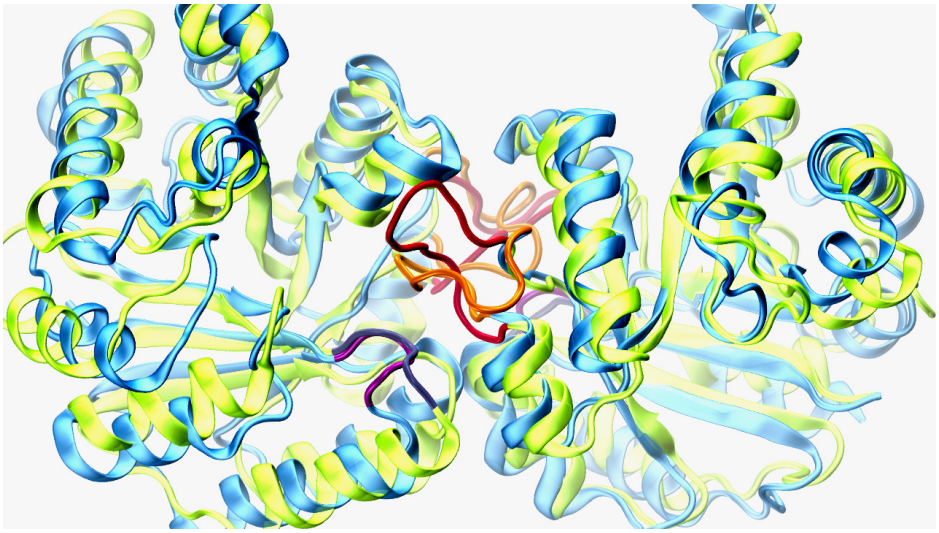

Figure 19. **Superimposition of the final structures (*t*=150ns) of reduced (cyan) and oxidized (green) PfTIM**, showing the conformational changes of loops 1 and 3 induced by the absence of cross-linking between mutant Tyr74Cys and Cys13. Loop 1 is shown in blue (reduced) and purple (oxidized), whereas loop 3 is shown in red (reduced) and orange (oxidized).

In terms of structural stability, reduced PfTIM showed an increase in the RMSD of 1 Å compared to oxidized PfTIM, indicating that cross-linking at the interface has a stabilizing effect. Backbone superimposition of oxidized and reduced PfTIM showed that these structural changes in the latter result from the loss of favorable contacts between loop 1 and 3 (**Figure 19**). We further calculated the side chain solvent-accessible surface

area (SASA) of Cys13 to evaluate to what extent cross-linking enforces solvent inaccessibility (**Figure 20**). Contrary to what it was expected, Cys13 of subunit A becomes more accessible to solvent in the oxidized form of PfTIM compared to its reduced form (**Figure 20A**). However, this pattern is different for subunit B, where SASA of Cys13 in the reduced form is larger than that observed in the oxidized PfTIM (Figure 20B). The exposure of Cys13 in both oxidized and reduced forms of PfTIM results in the decrease of intermolecular native contacts between loop 1 and loop 3 (**Figure 21**). Unlike the wild-type form of PfTIM, the oxidized PfTIM is unable to preserve more than 50% of the native contacts for more than a half of total simulation time; comparable results are obtained for reduced PfTIM.

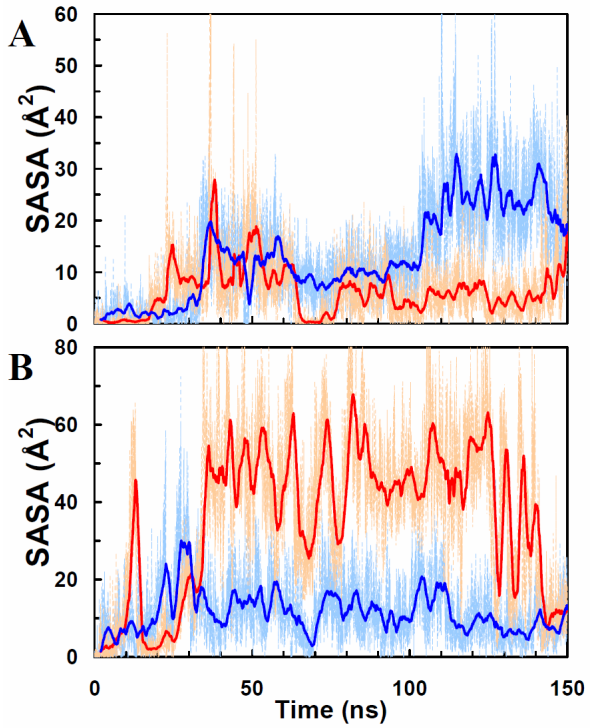

Figure 20. **Solvent-accessible surface area of Cys13 calculated for reduced (red line) and oxidized (blue line) PfTIM**. (A) and (B) correspond to Cys13 of subunits A and B, respectively.

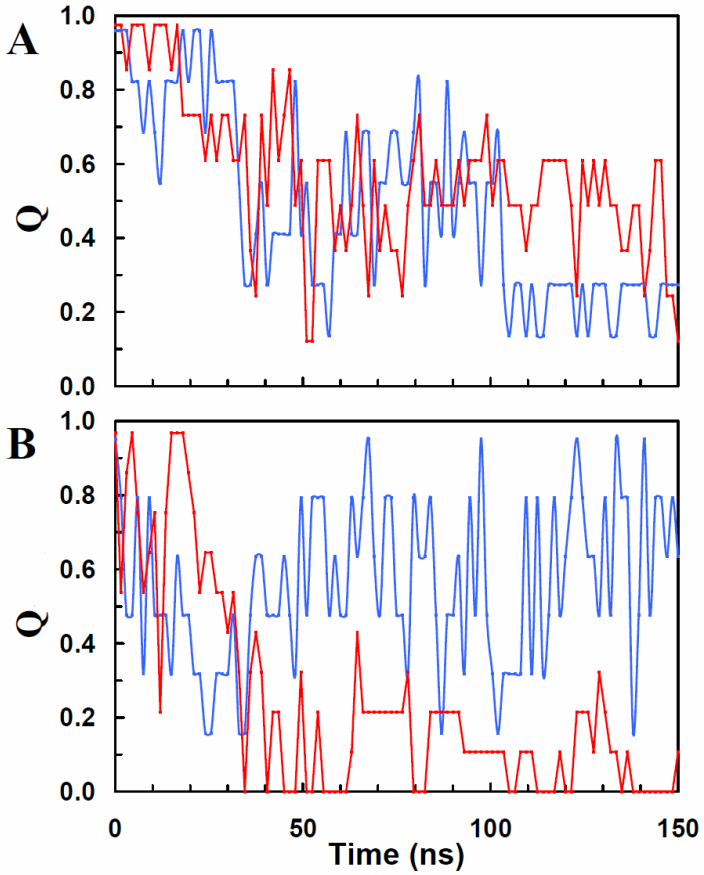

Figure 21. **Fraction of intermolecular native contacts between Cys13 and loop 3 computed for oxidized (blue line) and reduced (red line) PfTIM**. (A) Fraction of intermolecular contacts involving Cys13 of subunit A; (B) Fraction of intermolecular contacts involving Cys13 of subunit B.

In light of the fact that disulfide cross-linking neither reduces solvent accessibility of Cys13 nor restores the native contacts between loops 1 and 3, it is expected that the structure of loops 1 and 3 is dramatically affected in both oxidized and reduced forms of PfTIM. Superimposition of initial and final structures in the trajectories revealed that, while the native conformation of loop 3 is affected in both cross-linked and non-cross-linked dimers, the structure of loop 1 seems well preserved in the oxidized form. As the backbone conformation of Lys 12 is sensitive to large changes at the interface, we obtained the Ramachandran plots of this residue for each subunit of the oxidized and

reduced PfTIM (**Figure 22**). To our surprise, the native backbone conformation of Lys12 (φ≈50º, ψ≈-150º) does not change upon disulfide cross-linking (**Figure 22A**). In contrast, the absence of a disulfide bond between Cys13 and Cys74 produces a dramatic shift in the conformation of Lys12 (**Figure 22B**).

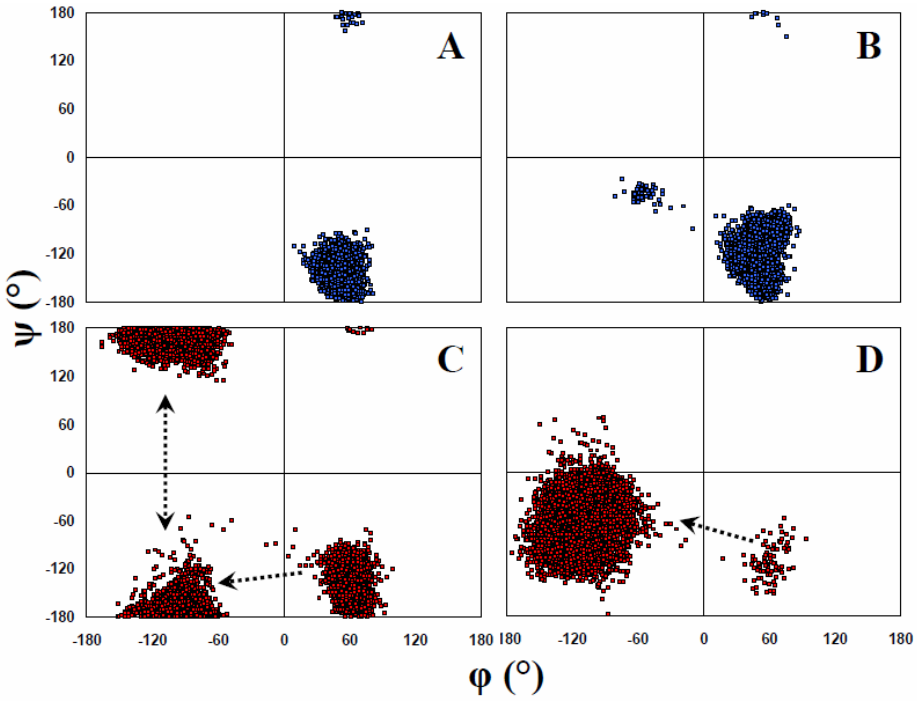

Figure 22. **Ramachandran plot of residue Lys12 in oxidized and reduced PfTIM**. (A) and (B) correspond to Lys12 of subunits A and B of oxidized PfTIM dimer, respectively. (C) and (D) correspond to Lys12 of subunits A and B of reduced PfTIM dimer, respectively.

3.6. Mobility of Tyr74 in the wild-type monomeric and dimeric PfTIM

We analyzed the evolutionary conservation of Tyr74 using the CONSURF server [92]. We observed that this position is structurally conserved, as its residue variety is only limited to tyrosine and phenylalanine. Considering that recent studies have suggested that the mobility of conserved residues in protein-protein interfaces is restricted [93], we assessed the flexibility of Tyr74 in both monomeric and dimeric PfTIM. Flexibility was

evaluated through side chain RMSD following the strategy of Rajamani et al. [94]. To avoid systematic errors caused by translation motions, we calculated side chain RMSD by aligning all heavy backbone atoms of Tyr74; RMSD was calculated for every single structure in the trajectories, and equilibrated structures were used as a reference.

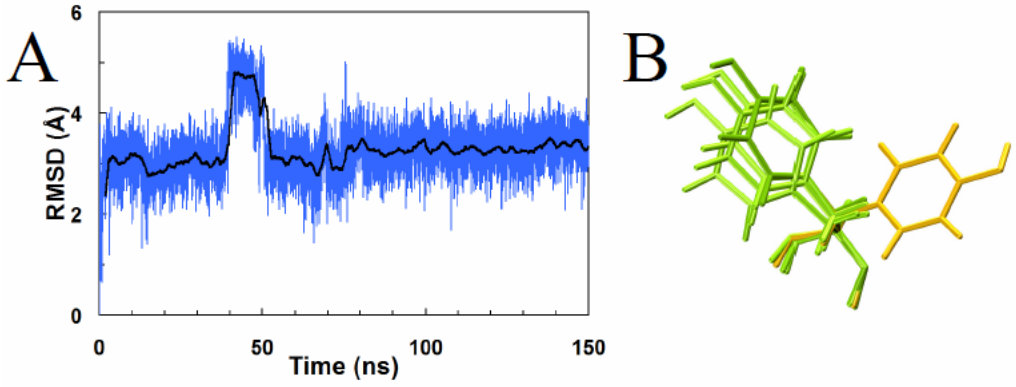

Figure 23. **(A) The root mean square deviation of Tyr74 of monomeric PfTIM with respect to the structure at *t*=0 ns**. (B) Representative conformations of Tyr74 extracted with the clustering procedure; green structures correspond to the most populated cluster found in the trajectory, whereas the green structure represents the conformation of Tyr74 between 40 and 50 ns.

Figure 23A shows that side chain RMSD of Tyr74 rapidly changes for about 3 Å in the first few nanoseconds of simulation of monomeric PfTIM. Except for the period of time between 40 and 50 ns, the side chain RMSD remains fairly steady throughout most of the simulation time, as expected for a structurally conserved residue. Most of the variations around the average RMSD (**Figure 23A**, black line) are the result thermal motions (i.e., the flipping of the aromatic ring in the picosecond timescale). Nonetheless, given the structural symmetry of the side chain, thermal motions have a negligible effect on the average mobility of Tyr74. This observation is supported by clustering using a maximum tolerance of 1 Å. This clustering procedure showed that Tyr74 side chain visits

two conformations during the entire simulation time; however, only one of these two conformations is significantly more populated (**Figure 23B**, green sticks).

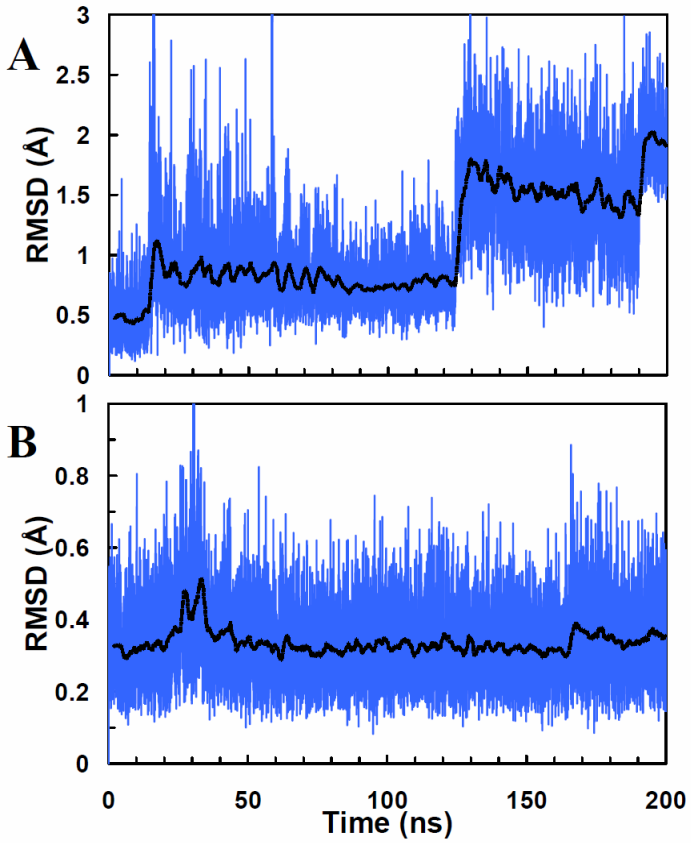

Figure 24. **The root mean square deviation of Tyr74 of dimeric PfTIM with respect to the structure at *t*=0 ns**. (A) corresponds to subunit A, whereas subunit B is labeled as (B).

Tyr74 of PfTIM dimer remains conformational restrained in the subunit A (**Figure 24A**). Three conformations are observed for this residue. The most frequent conformation, which resembles to that observed in the crystal structure, is present during the first 125 ns of simulation (~63% of the total simulation time). Although the second (~33% of the time) and third (~4% of the time) conformations of Tyr74 of subunit A deviate from the crystal structure, they appear only as a result of a slight repacking of the aromatic cluster Phe-69-Tyr74-Tyr101-Phe102. Tyr74 from subunit B remains very

constrained during the entire simulation time (<RMSD>≈ 0.5Å), indicating that its packing in the aromatic clusters is not altered in this timescale (**Figure 24B**).

3.7. Conservation of collective motions in the dimer

The motions of a protein may be broken down into their principal components by principal component analysis (PCA) [86-88]. By combining MD simulations and PCA, Cansu and Doruker showed that dimerization of chicken TIM gives rise to rigid-body motions that are crucial for enzymatic activity [95]. As a large number of critical interactions at the interface involve loops 1 and 3, it is expected that modifications in the geometry of this region will affect the native rigid-body motions of the dimer. Thus, we extracted the principal components from the wild-type, Tyr74Gly, oxidized and reduced forms of PfTIM.

The first principal component accounts for 37% and 40% of the overall motion in the wild-type and Tyr74Gly PfTIM dimer. For oxidized and reduced PfTIMs, the first principal component describes 44% and 29% of the overall motion of the dimer, respectively. We have plotted the rigid-body motions onto the three-dimensional structure of PfTIM (**Figure 25**). The rigid-body motion described by the first (and largest) principal component in wild-type PfTIM consists of a counter-rotation movement of the subunits (**Figure 25A**). This motion is consistent with that observed in the wild-type dimeric form of chicken TIM [95]. More importantly, this motion is strongly coupled to the opening/closing of active site loop 6 [95]. In Tyr74Gly dimeric PfTIM, this motion is still present (**Figure 25B**); however the homogeneity and magnitude of counter-rotation movement is affected by the removal of Tyr74 side chain. In the reduced

form of PfTIM dimer, where the removal of native contacts and the perturbation of the native conformation of Lys12 are more evident, the native rigid-body motions described by the first principal component are not longer present (**Figure 25D**). Surprisingly, the counter-rotation movement observed in wild-type PfTIM dimer appears to be conserved in the oxidized form of PfTIM (**Figure 25C**), despite the large loss of native contacts between loops 1 and 3 (**Figure 21**) and increased accessibility of Cys13 (**Figure 20**).

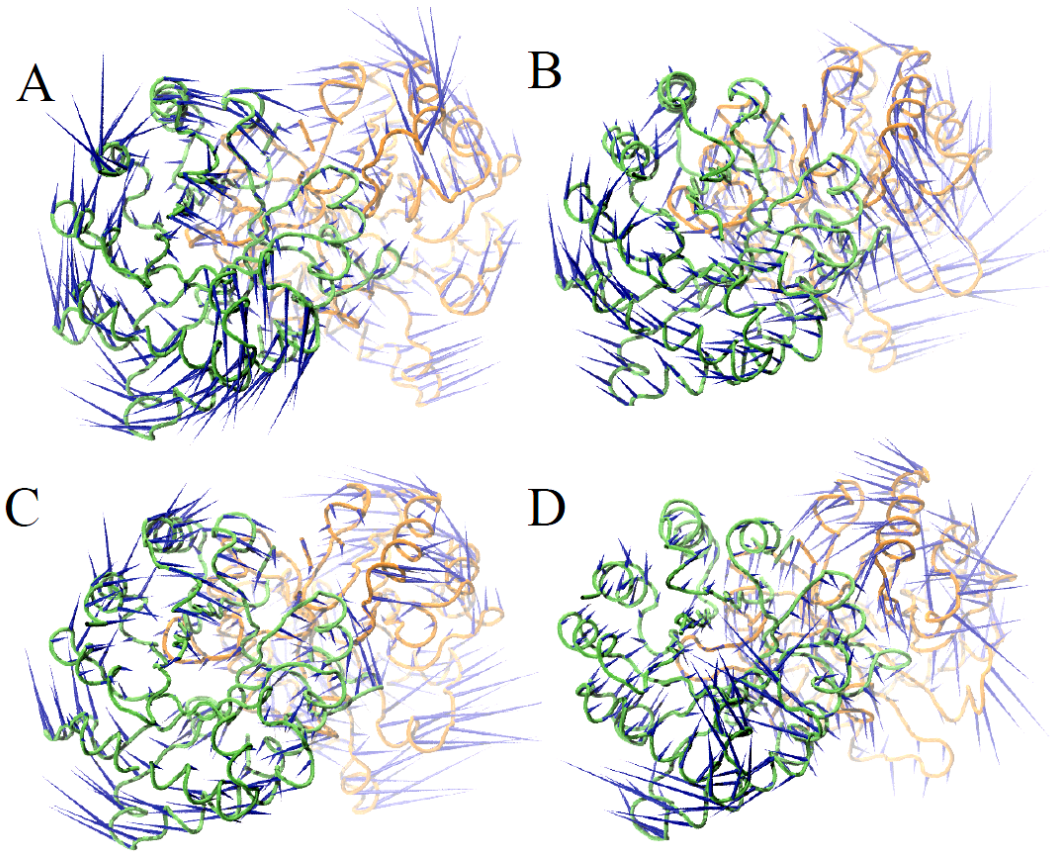

Figure 25. **Collective motions corresponding to the first principal component of (A) wild-type PfTIM, (B) Tyr74Gly PfTIM, (C) oxidized PfTIM and (D) reduced PfTIM**. The arrows, starting from the Cα of each residue, show the direction of the movement.

4. Discussion

4.1. Summary of Results

In the present study, we have performed MD simulations of PfTIM in the microsecond timescale to unveil the roles of the structurally conserved residue Tyr74 in the formation, stability and function of the PfTIM dimer. Here we studied two mutations of Tyr74 (Tyr74Gly and Tyr74Cys), provided that these mutants have been experimentally analyzed. Simulations of the monomeric PfTIM under physiological conditions revealed that mutant Tyr74Gly does not affect its folding and stability, in agreement with experimental data [35]. We also found that mutation of Tyr74 with glycine does not affect the native local dynamics of loop 3, which is essential for homodimerization. The trajectories of the dimer showed that the mutation does not affect the secondary and tertiary structure of its subunits; however, Tyr74Gly increases the flexibility of loops 1 and 3, decreases the fraction of native contacts between these two loops and alters the native bound-like backbone conformation of active site residue Lys12. Such effects were more noticeable when we used high temperature MD simulations to accelerate the mutant-induced conformational transitions of loops 1 and 3. Furthermore, the mutant also appears to affect the collective motions of the dimer that are required for catalytic activity. Disulfide cross-linking between Cys13 and the mutant Tyr74Cys appears to restore some of the native structural features of the interface (i.e. the bound-like backbone conformation of Lys12) and the wild-type-like collective motions of PfTIM dimer.

4.2. Tyr74 is a ready-made recognition motif for PfTIM homodimerization

In agreement with the experimental data obtained by Maithal et al. [35], we have shown that Tyr74Gly does not affect the folding of monomeric PfTIM. Unfortunately, these experimental studies did not provide any information regarding the local structural changes that this mutation might induce in the monomer. As Tyr74 is located at the tip of loop 3, it is possible that its mutation with glycine may (a) induce a local increase/decrease in flexibility of loop 3 or (b) produce long-range changes in other regions of the monomer via allosteric communication, which may directly affect the conformational entropy of association [96]. However, neither of these scenarios are the case here, as Tyr74Gly does not produce any noticeable shifts in the local dynamics of monomeric PfTIM (**Figure 11**). Yet, experiments have shown that monomeric PfTIM holding this mutation is unable to self-associate at low protein concentrations [35]. If Tyr74 does not have an effect in folding and stability of the dimer, how does it influence PfTIM dimerization?

Although a large percentage of monomeric TIM possesses a well-defined three-dimensional structure, approximately 20% of its primary structure remains natively disordered or unstructured. Recent experimental evidence have shown that intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs), which do not possess a well-defined 3D structure under physiological conditions are functional in their native state [97-100]. IDPs/IDRs are often found in protein-protein interactions (PPIs), thus fulfilling a number of key functional roles in the cell. Most importantly, many IDPs/IDRs undergo disorder-to-order transitions upon binding, ranging from the

restriction of the dynamics in the primary structure to the formation of a folded structure [101].

Here, we highlight the importance of intrinsic disorder in PfTIM dimerization as two of its regions, loops 1 and 3, are intrinsically disordered in the monomeric form of the enzyme. The intrinsically disordered nature of these loops was assessed by using the FoldUnfold server, which predicts whether a region of a protein in likely to be disordered or not [63, 102]. The crystal structure of PfTIM has shown that upon binding, both regions undergo a disorder-to-order transition, where loops 1 and 3 do not become structured but have an important reduction in their accessible conformational states. In our trajectories, we observed that the removal of Tyr74 causes a large increase in the dynamics of loops 1 and 3 and the subsequent loss of native-like contacts between these two IDRs; this result clearly indicates that Tyr74 plays an essential role in the binding-induced disorder-to-order transition of loops 1 and 3. This suggestion is supported by the fact that Tyr74 remains conformationally restricted even in the monomeric (unbound) form of PfTIM (**Figure 23**); conformationally restricted residues have been shown to act as anchor residues, which help avoiding kinetically costly structural rearrangements at the interface, allowing for a relatively smooth recognition process [94]. Here, the restricted mobility of the structurally conserved Tyr74 may facilitate the formation of native contacts between loops 1 and 3 without decreasing the binding kinetics via strong native-like contacts [103] and a 'fly-casting' mechanism arising from intrinsic disorder [104]. The rigidity of Tyr74 is also essential for specificity in molecular recognition as it provides a ready-made recognition motif [105] for PfTIM homodimerization via the formation of aromatic clusters at the interface. This recognition specificity may be

enhanced by other factors such as the type of aromatic-aromatic interaction (i.e., Phe-Phe or Phe-Tyr), the geometrical arrangement of the aromatic residues (i.e., stacked or T-shaped) and the hydrophobicity of the environment where the interaction takes place [58]. It is important to point out that, according to our simulations, Tyr74 may play a central role in a synergistic binding mechanism between loops 1 and 3. In this mechanism, both conformational selection and induced fit mechanisms participate in the binding of IDPs/IDRs, providing an efficient balance between kinetics and thermodynamics of binding [101]. In this particular case, the side chain of Tyr74 undergoes a conformational selection that produces an induced fit of loop 1, as the bound-like geometry of this loop is not natively populated in the monomer (**Figure 17**). This disorder-to-order transition is responsible for the formation of the particular geometrical arrangement of loop 1 that allows Cys13 to fully interact with loop 3 and Lys12 to be locked in a favorable position for catalysis.

4.3. The role of Tyr74 in the stability and function of PfTIM dimer

There is strong experimental evidence showing that Cys13 plays a crucial role in the stability of PfTIM dimer. For instance, carboxyl methylation of Cys13 induced the formation of monomeric species, as detected by gel filtration [60]. Furthermore, the mutant Cys13Asp exhibited a reduced stability to denaturants and 7-fold reduction in the enzymatic activity [60]. A similar effect was observed when Cys14 of *Trypanosoma brucei* TIM (equivalent to Cys13 in PfTIM) was either exposed to sulfhydryl reagents or replaced with the other nineteen amino acids. In the case of *Trypanosoma brucei* TIM, sulfhydryl reagents produced large structural changes and abolition of catalysis [106],

whereas mutations induced low stabilities and enzymatic activities compared to the wild-type dimer [107]. Although these studies have acknowledged the importance of Cys13 for the stability of the dimeric form of PfTIM, our simulations and experimental data suggest that Tyr74 plays an equally important role in the stability of PfTIM dimer.

What is the exact role that Tyr74 plays in the stability of the dimer? High-temperature MD simulations revealed that, once the side chain of Tyr74 is removed, Cys13-mediated favorable interactions are lost. This reduction of native contacts is the result of an increase in the accessible surface and a shift in the side chain orientation of Cys13. However, these simulations do not reveal whether this reduction of favorable interactions is the result of a cavity at the interface (as suggested by experiments [35, 36]) or a perturbation of the mechanical forces that stabilize Cys13-mediated interactions at the interface. In order to clarify the role of Tyr74 in the stability of the interactions at the interface, we analyzed the trajectories of oxidized and reduced PfTIM dimers. We observed that despite the presence of a cavity at the aromatic clusters of the interface, the reduction of native contacts and the increase of SASA of Cys13, the oxidized (cross-linked) form of PfTIM forces loop 1 to preserve its native-like geometry (**Figure 22A and B**). This native bound-like geometry in the dimer is characterized by positive Ramachandran φ values of Lys12 [24], which forces the side chain of Cys13 to interact favorably with loop 3 (**Figure 6**). In the absence of cross-linking, both native contacts and geometry of loop 1 are lost. These results show that disulfide cross-linking mimics the function of Tyr74 at the interface, indicating that Tyr74 possesses a mechanical role via aromatic-aromatic interactions with Phe69, Tyr101 and Phe102.

In addition to this mechanical role, Tyr74 actively participates in the function of PfTIM dimer. Recently, MD simulations of chicken TIM revealed that dimerization gives rise to collective motions that are absent in the monomer [95]. These collective motions, largely described by the first principal component, result in the counter-rotation of the subunits and the closure of loop 6. The results of these MD simulations suggested that these collective motions are necessary for TIM function [95]. Taking into consideration these observations, we extracted the collective motions described by the first principal component of wild-type, mutant, oxidized and reduced PfTIM dimer (**Figure 13**). Collective motions described by the first principal component of wild-type PfTIM are the same as for chicken TIM [95], suggesting that such motions are native in the dimeric TIM regardless of species. Surprisingly, Tyr74Gly mutant did not produce large structural changes at physiological temperature, but affected the native collective motions of the dimer (**Figure 25B**). What is even more surprising is that the oxidized form of PfTIM is able to maintain the native collective motions observed in the wild-type PfTIM (**Figure 25C**). Such motions are no longer present in the reduced form of PfTIM (**Figure 25D**). These results suggest that native collective motions that contribute to the catalytic efficiency of PfTIM strongly depend on the preservation of intermolecular interaction between Tyr74 and Cys13.

Hilser and Thompson have shown that site-to-site allosteric coupling is maximized when intrinsic disorder is present in the domains or segments containing one or both of the coupled binding sites [108]. Here, we observed that in the absence of Tyr74, the collective motions coupled to the closure motion of active site loop 6 are importantly affected. Hence, Tyr74 is possibly a link between intrinsic disorder and

allosterism in PfTIM, modulating subunit-subunit communication via coupled disorder-to-order transitions of loops 1 and 3, resulting in the collective motions observed in the dimeric form of PfTIM.

5. Conclusions

5.1. Tyr74 is not necessary for folding and structural stability of monomeric PfTIM.

5.2. Tyr74 acts as an anchor residue, which is necessary for monomer-monomer molecular recognition via aromatic interactions.

5.3. Tyr74 is required for the structural stability at the interface of the dimer.

5.4. Tyr74 participates in both the disorder-to-order transitions of loops 1 and 3 and the conformational stability of the catalytic residue Lys12.

5.5. Tyr74 facilitates the coupling of loops 1 and 3, which gives rise to collective motions necessary for the catalytic efficiency of PfTIM.

6. Perspectives

The first attempt to inhibit PfTIM through the disruption of monomer-monomer interactions was done by Singh et al. [61] Their approach consisted on using synthetic peptides whose sequences match loops 1 and 3. The authors observed that the peptide containing the sequence of loop 3 was able to inhibit dimerization with an IC_{50} of 600-600 μM. Although the inhibitory potency of this synthetic peptide is modest, the results showed that PfTIM can be inhibited via the disruption of the native contacts at the interface [61]. Similar approaches have been used to selectively inhibit trypanosomal TIM [64, 69]; however, in these cases small organic molecules were used instead of synthetic peptides, with surprisingly good results [65].

Our results showed a number of features of PfTIM that have not been taken into consideration in the rational design of drugs that perturb the stability of the dimer. First and foremost, the regions that account for most of the intermolecular contacts of dimeric PfTIM (loops 1 and 3) are intrinsically disordered. Upon binding, these regions undergo disorder-to-order transitions, which certainly play a determinant role in binding specificity. Furthermore, Tyr74 seems to modulate disorder-to-order transitions via aromatic-aromatic interactions at the interface. Considering this observations and previous drug design efforts, we propose that a synthetic molecule that inhibits PfTIM at the interface should fulfill the following characteristics: (a) it should be able to mimic Tyr74, necessary for binding at the aromatic clusters at the interface; that is, the pharmacophore model should include at least an aromatic moiety; (b) as intrinsic disorder is essential for PfTIM homodimerization, this feature should be mimicked by the potential inhibitor of PfTIM. Mimicking intrinsic disorder may increase binding

specificity for parasitic PfTIM without compromising effectiveness. Exploiting intrinsic disorder to target protein-protein interactions has been suggested recently [109, 110]; and (c) although the synthetic compound should not necessarily mimic the length of loop 3, it should be bulky enough to hinder the interaction between loops 1 and 3, which in turn will specifically prevent PfTIM homodimerization.

7. References

[1] D. L. Nelson, and M. M. Cox, Lehninger Principles of Biochemistry, Fourth ed. ed., W. H. Freeman, New York 2004.

[2] J. M. Herlihy, S. G. Maister, W. J. Albery, and J. R. Knowles, Energetics of triosephosphate isomerase: the fate of the 1(R)-3H label of tritiated dihydroxyacetone phsophate in the isomerase reaction, Biochemistry 15 (1976) 5601-5607.

[3] W. J. Albery, and J. R. Knowles, Free-energy profile of the reaction catalyzed by triosephosphate isomerase, Biochemistry 15 (1976) 5627-5631.

[4] W. J. Albery, and J. R. Knowles, Deuterium and tritium exchange in enzyme kinetics, Biochemistry 15 (1976) 5588-5600.

[5] R. C. Davenport, P. A. Bash, B. A. Seaton, M. Karplus, G. A. Petsko, and D. Ringe, Structure of the triosephosphate isomerase-phosphoglycolohydroxamate complex: an analogue of the intermediate on the reaction pathway, Biochemistry 30 (1991) 5821-5826.

[6] P. J. Lodi, and J. R. Knowles, Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications, Biochemistry 30 (1991) 6948-6956.

[7] T. K. Harris, R. N. Cole, F. I. Comer, and A. S. Mildvan, Proton transfer in the mechanism of triosephosphate isomerase, Biochemistry 37 (1998) 16828-16838.

[8] E. B. Nickbarg, R. C. Davenport, G. A. Petsko, and J. R. Knowles, Triosephosphate isomerase: removal of a putatively electrophilic histidine residue results in a subtle change in catalytic mechanism, Biochemistry 27 (1988) 5948-5960.

[9] D. L. Pompliano, A. Peyman, and J. R. Knowles, Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase, Biochemistry 29 (1990) 3186-3194.

[10] Z. Zhang, E. A. Komives, S. Sugio, S. C. Blacklow, N. Narayana, N. H. Xuong, A. M. Stock, G. A. Petsko, and D. Ringe, The role of water in the catalytic efficiency of triosephosphate isomerase, Biochemistry 38 (1999) 4389-4397.

[11] P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe, and M. Karplus, Computer simulation and analysis of the reaction pathway of triosephosphate isomerase, Biochemistry 30 (1991) 5826-5832.

[12] J. Aqvist, and M. Fothergill, Computer simulation of the triosephosphate isomerase catalyzed reaction, J Biol Chem 271 (1996) 10010-10016.

[13] G. Alagona, P. Desmeules, C. Ghio, and P. A. Kollman, Quantum mechanical and molecular mechanical studies on a model for the dihydroxyacetone phosphate-glyceraldehyde phosphate isomerization catalyzed by triose phosphate isomerase (TIM), J Am Chem Soc 106 (1984) 3623-3632.

[14] G. Alagona, C. Ghio, and P. A. Kollman, Simple model for the effect of Glu165 --> Asp165 mutation on the rate of catalysis in triose phosphate isomerase, J Mol Biol 191 (1986) 23-27.

[15] G. Alagona, C. Ghio, and P. A. Kollman, Do enzymes stabilize transition states by electrostatic interactions or pKa balance: The case of triose phosphate isomerase (TIM)?, J Am Chem Soc 117 (1995) 9855-9862.

[16] Q. Cui, and M. Karplus, Quantum mechanics/molecular mechanics studies of triosephosphate isomerase-catalyzed reactions: effect of geometry and tunneling on proton-transfer rate constants, J Am Chem Soc 124 (2002) 3093-3124.

[17] Q. Cui, and M. Karplus, Triosephosphate isomerase: a theoretical comparison of alternative pathways, J Am Chem Soc 123 (2001) 2284-2290.

[18] M. Perakyla, and T. A. Pakkanen, Ab initio models for receptor-ligand interactions in proteins. 4. Model assembly study of the catalytic mechanism of triosephosphate isomerase, Proteins 25 (1996) 225-236.

[19] T. K. Harris, C. Abeygunawardana, and A. S. Mildvan, NMR studies of the role of hydrogen bonding in the mechanism of triosephosphate isomerase, Biochemistry 36 (1997) 14661-14675.

[20] E. Maldonado, M. Soriano-Garcia, A. Moreno, N. Cabrera, G. Garza-Ramos, M. de Gomez-Puyou, A. Gomez-Puyou, and R. Perez-Montfort, Differences in the intersubunit contacts in triosephosphate isomerase from two closely related pathogenic trypanosomes, J Mol Biol 283 (1998) 193-203.

[21] M. E. Noble, R. K. Wierenga, A. M. Lambeir, F. R. Opperdoes, A. M. Thunnissen, K. H. Kalk, H. Groendijk, and W. G. Hol, The adaptability of the active site of trypanosomal triosephosphate isomerase as observed in the crystal structures of three different complexes, Proteins 10 (1991) 50-69.

[22] J. C. Williams, J. P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P. A. Michels, A. M. Lambeir, and R. K. Wierenga, Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power, Protein Eng 12 (1999) 243-250.

[23] A. Rodriguez-Romero, A. Hernandez-Santoyo, L. del Pozo Yauner, A. Kornhauser, and D. A. Fernandez-Velasco, Structure and inactivation of triosephosphate isomerase from Entamoeba histolytica, J Mol Biol 322 (2002) 669-675.

[24] S. S. Velanker, S. S. Ray, R. S. Gokhale, S. Suma, H. Balaram, P. Balaram, and M. R. Murthy, Triosephosphate isomerase from *Plasmodium falciparum*: the crystal structure provides insights into antimalarial drug design, Structure 5 (1997) 751-761.

[25] E. Lolis, T. Alber, R. C. Davenport, D. Rose, F. C. Hartman, and G. A. Petsko, Structure of yeast triosephosphate isomerase at 1.9-A resolution, Biochemistry 29 (1990) 6609-6618.

[26] Z. Zhang, S. Sugio, E. A. Komives, K. D. Liu, J. R. Knowles, G. A. Petsko, and D. Ringe, Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-A resolution, Biochemistry 33 (1994) 2830-2837.

[27] R. Aparicio, S. T. Ferreira, and I. Polikarpov, Closed conformation of the active site loop of rabbit muscle triosephosphate isomerase in the absence of substrate: evidence of conformational heterogeneity, J Mol Biol 334 (2003) 1023-1041.

[28] T. Kinoshita, R. Maruki, M. Warizaya, H. Nakajima, and S. Nishimura, Structure of a high-resolution crystal form of human triosephosphate isomerase: improvement of crystals using the gel-tube method, Acta Crystallogr Sect F Struct Biol Cryst Commun 61 (2005) 346-349.

[29] P. J. Lodi, L. C. Chang, J. R. Knowles, and E. A. Komives, Triosephosphate isomerase requires a positively charged active site: the role of lysine-12, Biochemistry 33 (1994) 2809-2814.

[30] M. E. Noble, J. P. Zeelen, and R. K. Wierenga, Structures of the "open" and "closed" state of trypanosomal triosephosphate isomerase, as observed in a new crystal form: implications for the reaction mechanism, Proteins 16 (1993) 311-326.

[31] J. C. Williams, and A. E. McDermott, Dynamics of the flexible loop of triosephosphate isomerase: the loop motion is not ligand gated, Biochemistry 34 (1995) 8309-8319.

[32] D. Joseph-McCarthy, E. Lolis, E. A. Komives, and G. A. Petsko, Crystal structure of the K12M/G15A triosephosphate isomerase double mutant and electrostatic analysis of the active site, Biochemistry 33 (1994) 2815-2823.

[33] E. Lolis, and G. A. Petsko, Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-A resolution: implications for catalysis, Biochemistry 29 (1990) 6619-6625.

[34] R. K. Wierenga, K. H. Kalk, and W. G. Hol, Structure determination of the glycosomal triosephosphate isomerase from Trypanosoma brucei brucei at 2.4 A resolution, J Mol Biol 198 (1987) 109-121.

[35] K. Maithal, G. Ravindra, G. Nagaraj, S. K. Singh, H. Balaram, and P. Balaram, Subunit interface mutation disrupting an aromatic cluster in *Plasmodium falciparum* triosephosphate isomerase: effect on dimer stability, Protein Eng 15 (2002) 575-584.

[36] B. Gopal, S. S. Ray, R. S. Gokhale, H. Balaram, M. R. Murthy, and P. Balaram, Cavity-creating mutation at the dimer interface of *Plasmodium falciparum* triosephosphate isomerase: restoration of stability by disulfide cross-linking of subunits, Biochemistry 38 (1999) 478-486.

[37] E. A. Meyer, R. K. Castellano, and F. Diederich, Interactions with aromatic rings in chemical and biological recognition, Angew Chem Int Ed Engl 42 (2003) 1210-1250.

[38] M. O. Sinnokrot, and C. D. Sherrill, Substituent effects in pi-pi interactions: sandwich and T-shaped configurations, J Am Chem Soc 126 (2004) 7690-7697.

[39] D. A. Evans, K. T. Chapman, D. T. Hung, and A. T. Kawaguchi, Transition state pi-solvation by aromatic rings: An electronic contribution to Diels-Alder reaction diastereoselectivity, Angew Chem Int Ed Engl 26 (1987) 1184-1185.

[40] A. V. Muehldorf, D. V. Engen, J. C. Warner, and A. D. Hamilton, Aromatic-aromatic interactions in molecular recognition: a family of artificial receptors for thymine that shows both face-to-face and edge-to-face orientations, J Am Chem Soc 110 (1988) 6561-6562.

[41] C. A. Hunter, Sequence-dependent DNA structure. The role of base stacking interactions, J Mol Biol 230 (1993) 1025-1054.

[42] P. Hobza, H. L. Selzle, and E. W. Schlag, Structure and properties of benzene-containing molecular clusters: nonempirical ab initio calculations and experiments, Chem Rev 94 (1994) 1767-1785.

[43] G. Kryger, I. Silman, and J. L. Sussman, Structure of acetylcholinesterase complexed with E2020 (Aricept): implications for the design of new anti-Alzheimer drugs, Structure 7 (1999) 297-307.

[44] H. Haviv, D. M. Wong, H. M. Greenblatt, P. R. Carlier, Y. P. Pang, I. Silman, and J. L. Sussman, Crystal packing mediates enantioselective ligand recognition at the peripheral site of acetylcholinesterase, J Am Chem Soc 127 (2005) 11029-11036.

[45] Y. Bourne, H. C. Kolb, Z. Radic, K. B. Sharpless, P. Taylor, and P. Marchot, Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation, Proc Natl Acad Sci U S A 101 (2004) 1449-1454.

[46] E. H. Rydberg, B. Brumshtein, H. M. Greenblatt, D. M. Wong, D. Shaya, L. D. Williams, P. R. Carlier, Y. P. Pang, I. Silman, and J. L. Sussman, Complexes of alkylene-linked tacrine dimers with Torpedo californica acetylcholinesterase: Binding of Bis5-tacrine produces a dramatic rearrangement in the active-site gorge, J Med Chem 49 (2006) 5491-5500.

[47] J. Correa-Basurto, J. Espinosa-Raya, M. Gonzalez-May, L. M. Espinoza-Fonseca, I. Vazquez-Alcantara, and J. Trujillo-Ferrara, Inhibition of acetylcholinesterase by two arylderivatives: 3a-Acetoxy-5H-pyrrolo (1,2-a) (3, 1)benzoxazin- 1,5-(3aH)-dione and cis-N-p-Acetoxy-phenylisomaleimide, J Enzyme Inhib Med Chem 21 (2006) 133-138.

[48] J. Correa-Basurto, C. Flores-Sandoval, J. Marin-Cruz, A. Rojo-Dominguez, L. M. Espinoza-Fonseca, and J. G. Trujillo-Ferrara, Docking and quantum mechanic studies on cholinesterases and their inhibitors, Eur J Med Chem 42 (2007) 10-19.

[49] R. Chelli, F. L. Gervasio, P. Procacci, and V. Schettino, Inter-residue and solvent-residue interactions in proteins: a statistical study on experimental structures, Proteins 55 (2004) 139-151.

[50] S. Marsili, R. Chelli, V. Schettino, and P. Procacci, Thermodynamics of stacking interactions in proteins, Phys Chem Chem Phys 10 (2008) 2673-2685.

[51] F. L. Gervasio, R. Chelli, P. Procacci, and V. Schettino, The nature of intermolecular interactions between aromatic amino acid residues, Proteins 48 (2002) 117-125.

[52] S. Aravinda, N. Shamala, C. Das, A. Sriranjini, I. L. Karle, and P. Balaram, Aromatic-aromatic interactions in crystal structures of helical peptide scaffolds containing projecting phenylalanine residues, J Am Chem Soc 125 (2003) 5308-5315.

[53] S. K. Burley, and G. A. Petsko, Aromatic-aromatic interaction: a mechanism of protein structure stabilization, Science 229 (1985) 23-28.

[54] S. K. Burley, and G. A. Petsko, Amino-aromatic interactions in proteins, FEBS Lett 203 (1986) 139-143.

[55] R. Mahalakshmi, S. Raghothama, and P. Balaram, NMR analysis of aromatic interactions in designed peptide beta-hairpins, J Am Chem Soc 128 (2006) 1125-1138.

[56] R. Mahalakshmi, G. Shanmugam, P. L. Polavarapu, and P. Balaram, Circular dichroism of designed peptide helices and beta-hairpins: analysis of Trp- and Tyr-rich peptides, Chembiochem 6 (2005) 2152-2158.

[57] G. B. McGaughey, M. Gagne, and A. K. Rappe, pi-Stacking interactions. Alive and well in proteins, J Biol Chem 273 (1998) 15458-15463.

[58] R. Chelli, F. L. Gervasio, P. Procacci, and V. Schettino, Stacking and T-shape competition in aromatic-aromatic amino acid interactions, J Am Chem Soc 124 (2002) 6133-6143.

[59] K. I. Cho, K. Lee, K. H. Lee, D. Kim, and D. Lee, Specificity of molecular interactions in transient protein-protein interaction interfaces, Proteins 65 (2006) 593-606.

[60] K. Maithal, G. Ravindra, H. Balaram, and P. Balaram, Inhibition of *Plasmodium falciparum* triose-phosphate isomerase by chemical modification of an interface cysteine.

Electrospray ionization mass spectrometric analysis of differential cysteine reactivities, J Biol Chem 277 (2002) 25106-25114.

[61] S. K. Singh, K. Maithal, H. Balaram, and P. Balaram, Synthetic peptides as inactivators of multimeric enzymes: inhibition of *Plasmodium falciparum* triosephosphate isomerase by interface peptides, FEBS Lett 501 (2001) 19-23.

[62] M. Banerjee, H. Balaram, and P. Balaram, Structural effects of a dimer interface mutation on catalytic activity of triosephosphate isomerase. The role of conserved residues and complementary mutations, Febs J 276 (2009) 4169-4183.

[63] O. V. Galzitskaya, S. O. Garbuzynskiy, and M. Y. Lobanov, FoldUnfold: web server for the prediction of disordered regions in protein chain, Bioinformatics 22 (2006) 2948-2949.

[64] V. Olivares-Illana, R. Perez-Montfort, F. Lopez-Calahorra, M. Costas, A. Rodriguez-Romero, M. Tuena de Gomez-Puyou, and A. Gomez Puyou, Structural differences in triosephosphate isomerase from different species and discovery of a multitrypanosomatid inhibitor, Biochemistry 45 (2006) 2556-2560.

[65] A. Tellez-Valencia, S. Avila-Rios, R. Perez-Montfort, A. Rodriguez-Romero, M. Tuena de Gomez-Puyou, F. Lopez-Calahorra, and A. Gomez-Puyou, Highly specific inactivation of triosephosphate isomerase from Trypanosoma cruzi, Biochem Biophys Res Commun 295 (2002) 958-963.

[66] L. M. Espinoza-Fonseca, and J. G. Trujillo-Ferrara, Exploring the possible binding sites at the interface of triosephosphate isomerase dimer as a potential target for anti-tripanosomal drug design, Bioorg Med Chem Lett 14 (2004) 3151-3154.

[67] L. M. Espinoza-Fonseca, and J. G. Trujillo-Ferrara, Structural considerations for the rational design of selective anti-trypanosomal agents: the role of the aromatic clusters at the interface of triosephosphate isomerase dimer, Biochem Biophys Res Commun 328 (2005) 922-928.

[68] L. M. Espinoza-Fonseca, and J. G. Trujillo-Ferrara, Toward a rational design of selective multi-trypanosomatid inhibitors: a computational docking study, Bioorg Med Chem Lett 16 (2006) 6288-6292.

[69] A. Tellez-Valencia, V. Olivares-Illana, A. Hernandez-Santoyo, R. Perez-Montfort, M. Costas, A. Rodriguez-Romero, F. Lopez-Calahorra, M. Tuena De Gomez-Puyou, and A. Gomez-Puyou, Inactivation of triosephosphate isomerase from Trypanosoma cruzi by an agent that perturbs its dimer interface, J Mol Biol 341 (2004) 1355-1365.

[70] W. F. van Gunsteren, and H. J. C. Berendsen, Algorithms for macromolecular dynamics and constraint dynamics, Mol Phys 34 (1977) 1311-1327.

[71] W. Weber, P. H. Нънenberger, and J. A. McCammon, Molecular Dynamics Simulations of a Polyalanine Octapeptide under Ewald Boundary Conditions: Influence of Artificial Periodicity on Peptide Conformation, J Phys Chem B 104 (2000) 3668-3675.

[72] M. P. Allen, and D. J. Tildesley, Computer simulation of liquids, ed., Oxford University Press, Oxford 1987.

[73] T. Darden, D. York, and L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems., J Chem Phys 98 (1993) 10089-10092.

[74] U. Essmann, L. Perera, and M. L. Berkowitz, A smooth particle mesh Ewald method, J Chem Phys 103 (1995) 8577-8593.

[75] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. A. Kollman, AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, Comput Phys Commun 91 (1995).

[76] J. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. Reiher, W.E. , B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J Phys Chem B 102 (1998) 3586-3616.

[77] W. L. Jorgensen, and J. Tirado_Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, J Am Chem Soc 110 (1988) 1657-1666.

[78] A. D. MacKerell, Jr., M. Feig, and C. L. Brooks, 3rd, Improved treatment of the protein backbone in empirical force fields, J Am Chem Soc 126 (2004) 698-699.

[79] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, Scalable molecular dynamics with NAMD, J Comput Chem 26 (2005) 1781-1802.

[80] H. Li, A. D. Robertson, and J. H. Jensen, Very fast empirical prediction and rationalization of protein pKa values, Proteins 61 (2005) 704-721.

[81] D. Frishman, and P. Argos, Knowledge-based protein secondary structure assignment, Proteins 23 (1995) 566-579.

[82] W. Kabsch, and C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577-2637.

[83] W. Humphrey, A. Dalke, and K. Schulten, VMD: visual molecular dynamics, J Mol Graph 14 (1996) 33-38, 27-38.

[84] E. A. Coutsias, C. Seok, and K. A. Dill, Using quaternions to calculate RMSD, J Comput Chem 25 (2004) 1849-1857.

[85] W. Kabsch, A solution for the best rotation to relate two sets of vectors, Acta Crystallogr 32 (1976) 922-923.

[86] A. E. Garcia, Large-amplitude nonlinear motions in proteins, Phys Rev Lett 68 (1992) 2696-2699.

[87] M. A. Balsera, W. Wriggers, Y. Oono, and K. Schulten, Principal component analysis and long time protein dynamics, J Phys Chem 100 (1996) 2567-2572.

[88] A. Amadei, A. B. Linssen, and H. J. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412-425.

[89] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, GROMACS: fast, flexible, and free, J Comput Chem 26 (2005) 1701-1718.

[90] C. P. Barrett, B. A. Hall, and M. E. Noble, Dynamite: a simple way to gain insight into protein motions, Acta Crystallogr D Biol Crystallogr 60 (2004) 2280-2287.

[91] P. A. Karplus, Experimentally observed conformation-dependent geometry and hidden strain in proteins, Protein Sci 5 (1996) 1406-1420.

[92] M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko, and N. Ben-Tal, ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, Nucleic Acids Res 33 (2005) W299-302.

[93] O. N. Yogurtcu, S. B. Erdemli, R. Nussinov, M. Turkay, and O. Keskin, Restricted mobility of conserved residues in protein-protein interfaces in molecular simulations, Biophys J 94 (2008) 3475-3485.

[94] D. Rajamani, S. Thiel, S. Vajda, and C. J. Camacho, Anchor residues in protein-protein interactions, Proc Natl Acad Sci U S A 101 (2004) 11287-11292.

[95] S. Cansu, and P. Doruker, Dimerization affects collective dynamics of triosephosphate isomerase, Biochemistry 47 (2008) 1358-1368.

[96] K. K. Frederick, M. S. Marlow, K. G. Valentine, and A. J. Wand, Conformational entropy in molecular recognition by proteins, Nature 448 (2007) 325-329.

[97] P. E. Wright, and H. J. Dyson, Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm, J Mol Biol 293 (1999) 321-331.

[98] A. K. Dunker, and Z. Obradovic, The protein trinity--linking function and disorder, Nat Biotechnol 19 (2001) 805-806.

[99] P. Tompa, Intrinsically unstructured proteins, Trends Biochem Sci 27 (2002) 527-533.

[100] V. N. Uversky, Natively unfolded proteins: a point where biology waits for physics, Protein Sci 11 (2002) 739-756.

[101] L. M. Espinoza-Fonseca, Reconciling binding mechanisms of intrinsically disordered proteins, Biochem Biophys Res Commun 382 (2009) 479-482.

[102] S. O. Garbuzynskiy, M. Y. Lobanov, and O. V. Galzitskaya, To be folded or to be unfolded?, Protein Sci 13 (2004) 2871-2877.

[103] L. M. Espinoza-Fonseca, Thermodynamic aspects of coupled binding and folding of an intrinsically disordered protein: a computational alanine scanning study, Biochemistry 48 (2009) 11332-11334.

[104] B. A. Shoemaker, J. J. Portman, and P. G. Wolynes, Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, Proc Natl Acad Sci U S A 97 (2000) 8868-8873.

[105] S. R. Kimura, R. C. Brower, S. Vajda, and C. J. Camacho, Dynamical view of the positions of key side chains in protein-protein recognition, Biophys J 80 (2001) 635-642.

[106] G. Garza-Ramos, N. Cabrera, E. Saavedra-Lira, M. Tuena de Gomez-Puyou, P. Ostoa-Saloma, R. Perez-Montfort, and A. Gomez-Puyou, Sulfhydryl reagent susceptibility in proteins with high sequence similarity--triosephosphate isomerase from Trypanosoma brucei, Trypanosoma cruzi and Leishmania mexicana, Eur J Biochem 253 (1998) 684-691.

[107] G. Hernandez-Alcantara, G. Garza-Ramos, G. M. Hernandez, A. Gomez-Puyou, and R. Perez-Montfort, Catalysis and stability of triosephosphate isomerase from Trypanosoma brucei with different residues at position 14 of the dimer interface. Characterization of a catalytically competent monomeric enzyme, Biochemistry 41 (2002) 4230-4238.

[108] V. J. Hilser, and E. B. Thompson, Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins, Proc Natl Acad Sci U S A 104 (2007) 8311-8315.

[109] R. B. Russell, and T. J. Gibson, A careful disorderliness in the proteome: sites for interaction and targets for future therapies, FEBS Lett 582 (2008) 1271-1275.

[110] Y. Cheng, T. LeGall, C. J. Oldfield, J. P. Mueller, Y. Y. Van, P. Romero, M. S. Cortese, V. N. Uversky, and A. K. Dunker, Rational drug design via intrinsically disordered protein, Trends Biotechnol 24 (2006) 435-442.

Archives of Biochemistry and Biophysics xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Tyr74 is essential for the formation, stability and function of *Plasmodium falciparum* triosephosphate isomerase dimer

L. Michel Espinoza-Fonseca [a,b,*], Carlos Wong-Ramírez [b], José G. Trujillo-Ferrara [c]

[a] *Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA*
[b] *Departamento de Bioquímica, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Prol. Carpio y Plan de Ayala, Mexico City 11340, Mexico*
[c] *Departamento de Bioquímica, Escuela Superior de Medicina, Instituto Politécnico Nacional, Plan de San Luis y Díaz Mirón, Mexico City 11340, Mexico*

ARTICLE INFO

Article history:
Received 16 September 2009
and in revised form 9 November 2009
Available online xxxx

Keywords:
Triosephosphate isomerase
Molecular dynamics simulations
Intrinsically disordered proteins/regions
Disorder-to-order transitions
Aromatic–aromatic interactions

ABSTRACT

Plasmodium falciparum triosephosphate isomerase (PfTIM) is known to be functional only as a homodimer. Although many studies have shown that the interface Cys13 plays a major role in the stability of the dimer, a few reports have demonstrated that structurally conserved Tyr74 may be essential for the stability of PfTIM dimer. To understand the role of Tyr74, we have performed molecular dynamics (MD) simulations of monomeric and dimeric PfTIM mutated to glycine and cysteine at position 74. Simulations of the monomer revealed that mutant Tyr74Gly does not produce changes in folding and stability of the monomer. Interestingly, comparison of the flexibility of Tyr74 in the monomer and dimer revealed that this residue possesses an intrinsic restricted mobility, indicating that Tyr74 is an anchor residue required for homodimerization. Tyr74 also appears to play an important role in binding by facilitating the disorder-to-order transitions of loops 1 and 3, which allows Cys13 to form favorable interactions with loop 3 and Lys12 to be locked in a favorable position for catalysis. High-temperature MD simulations of the wild-type and Tyr74Gly PfTIM dimers showed that the aromatic moiety of Tyr74 is necessary to preserve the geometry and native contacts between loops 1 and 3 at the interface of the dimer. Disulfide cross-linking between mutant Tyr74Cys and Cys13 further revealed that Tyr74 stabilizes the geometry of loop 1 (which contains the catalytic residue Lys12) and the interactions between loops 1 and 3 via aromatic–aromatic interactions with residues Phe69, Tyr101, and Phe102. Principal component analysis showed that Tyr74 is also necessary to preserve the collective motions in the dimer that contribute to the catalytic efficiency of PfTIM dimer. We conclude that Tyr74 not only plays a role in the stability of the dimer, but also participates in the dimerization process and collective motions via coupled disorder-to-order transitions of intrinsically disordered regions, necessary for efficiency in the catalytic function of PfTIM.

Introduction

Triosephosphate isomerase (TIM)[1] is an enzyme which plays a key role in the first phase of the glycolytic pathway as it catalyzes the rapid and reversible isomerization of dihydroxyacetone phosphate to glyceraldehyde-3-phosphate. Additionally, TIM also has an important role in gluconeogenesis, the hexosemonophosphate shunt and fatty acid biosynthesis. TIM has been labeled as a "perfect enzyme" because the chemical steps of the reaction are accelerated sufficiently so that the diffusion steps are rate-limiting [1–3].

Like most TIMs, *Plasmodium falciparum* TIM (PfTIM) is functional only as a homodimer. Each monomer composed of 247 residues folds as a classical α/β-barrel structure, where eight β-sheets and eight α-helices form the inner and outer layers of the barrel, respectively; these α-helices and β-sheets are connected to each other by loops [4]. The highly conserved active site of PfTIM is located close to the C-terminus of the barrel, as it has been found in all other TIMs.

The overall surface area buried in the interface of PfTIM is ~1800 Å^2 per subunit, which constitutes 15.5% of the total solvent-accessible area of the isolated subunits. In the PfTIM dimer, 12% of the total polar surface and 19% of the total non-polar surface becomes solvent inaccessible upon dimerization [4]. The largest contribution to this buried area comes from loop 3, which interacts with loop 1 (particularly with residue Cys13) of the other subunit (Fig. 1A). Loop 3 (residues 69–79), which protrudes ~13 Å out of the bulk of the monomer and docks to a narrow pocket close to the active site of the other subunit, contributes substantially to

* Corresponding author. Address: Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA.
E-mail address: mef@ddt.biochem.umn.edu (L.M. Espinoza-Fonseca).

[1] *Abbreviations used:* PfTIM, *Plasmodium falciparum* triosephosphate isomerase; MD, molecular dynamics; TIM, triosephosphate isomerase; VMD, visual molecular dynamics; PCA, principal component analysis; RMSD, root-mean-square deviation; RMSF, root-mean-square fluctuations.

0003-9861/$ - see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.abb.2009.11.009

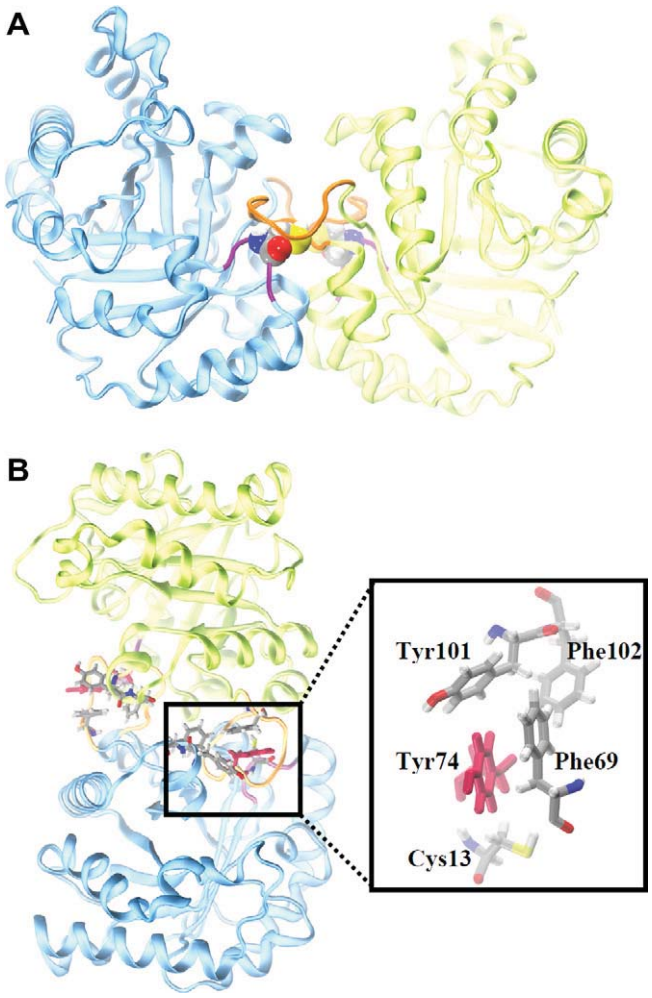

Fig. 1. (A) X-ray structure of dimeric PfTIM. Loops 1 and 3 are colored as purple and orange, respectively; interface Cys13 is shown as van der Waals spheres. (B) Location of the Tyr74-mediated aromatic clusters at the interface of PfTIM; the square shows an enlarged view of these aromatic clusters. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

intersubunit interactions [4]. This loop is crucial for dimer stability as ~80% of the intersubunit atom–atom contacts are involved in this region, a feature noted in earlier studies of other TIMs [5–7].

Analysis of the 3D structure of PfTIM has revealed that dimer stability arises from a combination of polar and non-polar interactions. A large amount of interactions occur between Cys13 (loop 1) from one subunit and loop 3 from the other subunit, forming a solvent-inaccessible polar cage. In addition to Cys13, this polar cage is constituted by the backbone of Phe69, Asn71, Gly72, Ser73, Tyr74 and Glu77; side chain–side chain interactions are also observed between Cys13 and Ser67, Glu77 and Ser79. Polar interactions around this region have been also observed; for instance, a subunit–subunit salt bridge between Glu77 and Arg98 was found in the X-ray structure of PfTIM [4]. Among non-polar interactions that stabilize the dimer, a relatively large hydrophobic patch is observed at the interface of TIM. Such hydrophobic patch is formed by residues Leu17, Val44 and Val46 of one subunit and Ile63, Val78, Ile82, Ala83, Leu86 and Ile88 of the other subunit.

Although experimental studies have acknowledged the importance of residue Cys13 [8] as well as loops 1 and 3 [9] for the stability of PfTIM, the role of the structurally conserved interface residue Tyr74 in the stability and function of TIM has not been studied extensively. The backbone of Tyr74 interacts with Cys13 via polar interactions, whereas its side chain participates in two symmetrical aromatic clusters at the interface (Fig. 1B). These aromatic clusters are formed by Phe69 and Tyr74 from the loop 3 of one subunit and Tyr101 and Phe102 from the adjacent subunit (Fig. 1B). Gopal et al. studied the mutation of Tyr74Cys with the aim of creating a stable PfTIM dimer via disulfide cross-linking with interface Cys13. Surprisingly, they observed that the oxidized form (cross-linked) of TIM was significantly more stable than its reduced (non-crossed-linked) counterpart, as revealed by thermal precipitation studies [10]. Following up with this study, Maithal et al. engineered the mutation Tyr74Gly in order to determine to what extent the stability of PfTIM dimer is affected by a perturbation of the interface aromatic clusters [11]. The results revealed that the stability and enzymatic activity of the mutant were significantly reduced. Fluorescence and circular dichroism experiments showed that both wild-type and Tyr74Gly PfTIMs displayed similar spectroscopic properties, suggesting that the mutation does not affect folding [11]. Further gel filtration of mutant PfTIM revealed that monomeric and dimeric species are in dynamic equilibrium, with the former predominating at low protein concentration. Low concentrations of urea (<2 M) were also found to drive the equilibrium of mutant PfTIM toward its monomeric form [11]. More recently, the triple mutant Trp11Phe/Trp168Phe/Tyr74Trp was studied; this mutant possessed a 20-fold reduction in activity compared to the wild-type enzyme [12].

While these studies have provided some valuable information regarding the role of Tyr74 in the stability of the PfTIM dimer, its involvement in the formation and enzymatic function at atomic level remains unknown. In this study, we applied classical molecular dynamics (MD) simulations (cumulative total simulation time of 1.3 μs) to study the role of Tyr74 in the stability and dynamics of dimeric PfTIM. In accordance with experimentally studied mutations of Tyr74 [10,11], we performed eight MD simulations of PfTIM: (a) monomeric and dimeric wild-type at 310 K, (b) monomeric and dimeric mutant (Tyr74Gly) at 310 K, (c) dimeric wild-type and mutant (Tyr47Gly) at 400 K, (d) mutant (Tyr74Cys) at 333 K and (e) mutant (Tyr74Cys) cross-linked with Cys13. The results not only support experimental data on the stabilizing effect of Tyr74, but also revealed its participation in the dimerization process and collective motions via coupled disorder-to-order transitions of intrinsically disordered regions that give rise to the catalytic function of PfTIM.

Computational methods

Preparation of monomeric PfTIM

The crystal structure of PfTIM (resolution of 2.2 Å) was used as starting coordinates for the simulations (PDB accession code: 1YDV). Considering that the native structure of PfTIM is homodimer in the crystal structure, only the subunit A (as labeled is the crystal structure) was used in this study to simulate the dynamics of monomeric PfTIM. The rationale for choosing subunit A is that there are no structural differences between the two monomers, as revealed by backbone superimposition. The Tyr74Gly mutation was modeled using PSFGEN, a standalone program included in NAMD 2.6 [13]. Side-chain ionization states of both monomeric wild-type and Tyr74Gly PfTIMs were adjusted to a pH of 7.0 using the program PROPKA [14]. Hydrogen atoms and $-NH_3^+$ and $-COO^-$ termini were added using PSFGEN. Both monomers were placed in the center of a TIP3P water box with a margin of ~20 Å between the protein and the boundaries of the periodic box. Chlorine and sodium counterions were added to yield a neutral charge on the system, and to produce a physiological ionic strength. Protein, water and ions were modeled with the CHARMM 27 force field [15,16].

Please cite this article in press as: L.M. Espinoza-Fonseca et al., Arch. Biochem. Biophys. (2009), doi:10.1016/j.abb.2009.11.009

Table 1
Summary of the characteristics and conditions of the systems studied.

System	Number of atoms	Number of waters	Temperature (K)	Ensemble used	Length of simulation (ns)
Wild-type PfTIM, monomer	75,023	23,681	310	NPT	150
Tyr74Gly PfTIM, monomer	75,024	23,686	310	NPT	150
Wild-type PfTIM, dimer	120,546	37,544	310	NPT	200
Tyr74Gly PfTIM, dimer	120,548	37,548	310	NPT	200
Wild-type PfTIM, dimer	120,546	37,544	400	NVT	150
Tyr74Gly PfTIM, dimer	120,548	37,548	400	NVT	150
Tyr74Cys PfTIM, oxidized	120,186	37,426	333	NPT	150
Tyr74Cys PfTIM, reduced	120,235	37,441	333	NPT	150

Preparation of dimeric PfTIM

As we are interested in the native dynamics of the monomer and its perturbation by the mutation of Tyr74 to glycine, we prepared the wild-type and mutated PfTIM dimer. Considering that the native structure of the triosephosphate isomerase is a homodimer, two identical mutations at the interface (i.e., one per monomer) were modeled at the interface of PfTIM. The structure of the wild-type PfTIM was retrieved from the Protein Data Base (PDB code: 1YDV). The mutation Tyr74Gly was generated with PSFGEN. Side-chain ionization states of the dimers were adjusted to a pH of 7.0 using the program PROPKA [14]. N and C termini were kept as $-NH_3^+$ and $-COO^-$, respectively. Wild-type and mutant dimers were placed in a box of TIP3P water with a margin of ~20 Å between the protein and the boundaries of the periodic box. Chlorine and sodium counterions were added to produce a neutral charge on the system, and to produce a ionic strength of ~150 mM. Protein, water and ions were modeled with the CHARMM 27 force field [15,16].

Preparation of disulfide cross-linked and non-cross-liked PfTIM

In accordance with experimental studies [10], we modeled the symmetric mutant Tyr74Cys using the procedure described above. Two individual systems were prepared: one where each Tyr74Cys mutant at the interface is left in its reduced form and one where Cys74 was cross-linked with Cys13 from the adjacent subunit (oxidized form). The ionization states of each dimer were assigned using PROPKA [14], and hydrogen atoms were added with PSFGEN. Each system was embedded in a box of water with margins of ~20 Å. The ionic strength of the systems was further adjusted to ~150 mM.

Molecular dynamics simulations

Molecular dynamics simulations were performed using the program NAMD 2.6 [13]. An NPT ensemble was used for systems studied at 310 K and 333 K, whereas an NVT ensemble was used for systems simulated at a temperature of 400 K; in the latter case, the volume of the systems was adjusted to that obtained at 310 K, and fixed throughout the simulation time. Periodic boundary conditions [17] were imposed on the systems. The electrostatic term is described by using the particle mesh-Ewald algorithm [18,19]. The non-bonded cutoff, switching distance and non-bonded pair-list distance will be set to 9, 8 and 10.5 Å, respectively. The SHAKE [20] algorithm for bonds to hydrogen atoms allowed a 2 fs time step; the impulse-based Verlet-I multi-step method was used with 2 fs for bonded, 4 fs for non-bonded and 8 fs for long-range electrostatic forces. Constant pressure (for systems simulated under constant pressure conditions) and temperature on the systems are maintained with an isotropic Langevin barostat and a Langevin thermostat, respectively. Thousand steps of conjugate gradient algorithm were used to minimize each system with restraints to protein backbone, followed by 1000 steps without restraints. Systems were warmed up for 60 ps each and equilibrated for 2 ns with lower restraints. Production runs were continued for eight independent systems as described in Table 1, for a cumulative total simulation time of 1.3 μs.

Analysis of trajectories

The Visual Molecular Dynamics (VMD [21]) program was used for visualization and analysis of trajectories. Principal component analysis (PCA) of the trajectories was performed using GROMACS [22], and generation of the plots describing collective motions was done using the Dinatraj server [23].

Results

Structural stability of monomeric PfTIM

An important feature we address in this study is whether Tyr74Gly mutation can induce considerable perturbation in the native secondary structure content of monomeric PfTIM. To determine the effect of the mutant on the secondary structure, we calculated the fraction of secondary structure averaged over the 150 ns of simulation (Fig. S1, Supplementary material). We found that the structural fractions are almost identical for both wild-type and mutant monomeric PfTIM, clearly indicating that mutation of Tyr74 to glycine does not produce any significant changes in the stability of the secondary structure of the monomer, in agreement with circular dichroism spectroscopy [11].

In order to evaluate the effect of Tyr74Gly on the structural stability and global dynamics of the monomer, we computed root-mean-square deviation (RMSD) of each structure in the trajectories of the wild-type and mutant PfTIMs. RMSD was calculated by superimposing the Cα trace of each snapshot using the structure at $t = 0$ ns as a reference (Fig. S2, Supplementary material). RMSD plots showed that both monomers have not extensively deviated from the reference structure, implying that the tertiary structure is largely conserved despite the mutation of Tyr74 to glycine. In addition, RMSD values do not differ much between monomeric wild-type and mutant PfTIM in the 150-ns trajectories, suggesting that in this timescale the mutant Tyr74Gly does not affect the global dynamics of the monomer.

Root-mean-square fluctuations (RMSF) of Cα atoms about their average positions were computed for each residue of wild-type and mutant PfTIM monomers (Fig. 2). The most mobile parts of wild-type monomeric PfTIM are the connecting loops and the N- and C-termini of the protein. Loops 1, 3 and 6 are especially more flexible compared to the rest of the protein; such increased flexibility is not surprising, as loops 1 and 3 undergo disorder-to-order transitions upon dimerization, whereas the intrinsic flexibility of loop 6 is essential for catalysis [24,25]. Although replacement of Tyr74 by glycine did not produce any significant change in the secondary and tertiary structure of monomeric PfTIM, the mutation might

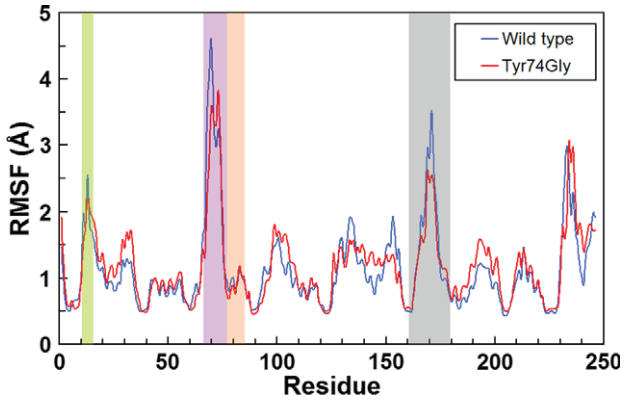

Fig. 2. Root-mean-square fluctuations of Cα atoms about their average positions computed for wild-type and Tyr74Gly PfTIM monomers. Green, purple and gray shading indicate the location of loops 1, 3 and 6, respectively. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

produce noticeable changes in the local dynamics of loop 3 (Fig. 2). Interestingly, Tyr74Gly did not induce important changes in the local flexibility, as the average values of RMSF of loop 3 are 2.45 Å and 2.25 Å for wild-type and mutant monomers, respectively. Likewise, Tyr74Gly did not seem to appreciably affect the mobility of Cα atoms of other regions compared to the wild-type monomer.

Effect of the mutation Tyr74Gly on the structure of the dimer

Experiments have shown that Tyr74Gly significantly reduces the stability of the dimer [11], although the exact destabilization mechanism is not yet clear. Therefore, we performed MD simulations of wild-type and Tyr74Gly dimers. Analysis of the fraction of secondary structure showed that mutation of Tyr74 to glycine does not significant affect the folding of the subunits of dimeric PfTIM (Table 2). Likewise, the tertiary structure of each subunit remained unchanged during the 200 ns of simulation. Time-dependent RMSD was calculated for wild-type and mutant dimers by superimposing the Cα trace of each snapshot using the coordinates of equilibrated dimers as a reference (Fig. S3, Supplementary material). RMSD of the wild-type dimer did not largely deviate from the original structure (<RMSD> = 2 Å) in this timescale. Similarly, the quaternary structure of the mutated dimer remained fairly stable between 0 and 140 ns (<RMSD> = 2 Å); however, a sudden increase of 1 Å in the RMSD was observed in the mutant at t = 145 ns.

To determine if this increase in the RMSD is the result of a conformational perturbation at the interface due to the removal of Tyr74, we calculated the RMSF of each residue and subunit of wild-type and mutant dimers (Fig. 3). Compared to the monomeric form of PfTIM (Fig. 2), loops 1 and 3 of each subunit of the wild-type dimer are considerably less flexible, although active site loop 6 conserves its flexible nature. These structural properties inherent to the wild-type form of the dimer are not observed when the mutation Tyr74Gly is introduced at the interface. This mutation increases the flexibility of loop 1 in both subunits compared to the wild-type

Table 2
Percentage of secondary structure content for wild-type and mutant dimers at 310 K.

	Wild-type (Monomer A)	Wild-type (Monomer B)	Tyr74Gly (Monomer A)	Tyr74Gly (Monomer B)
α-Helix	45 ± 2	46 ± 2	45 ± 2	45 ± 2
β-Sheet	18 ± 1	17 ± 1	17 ± 1	17 ± 1
Turn	18 ± 3	17 ± 2	18 ± 3	17 ± 2
Coil	19 ± 2	20 ± 2	20 ± 2	21 ± 2

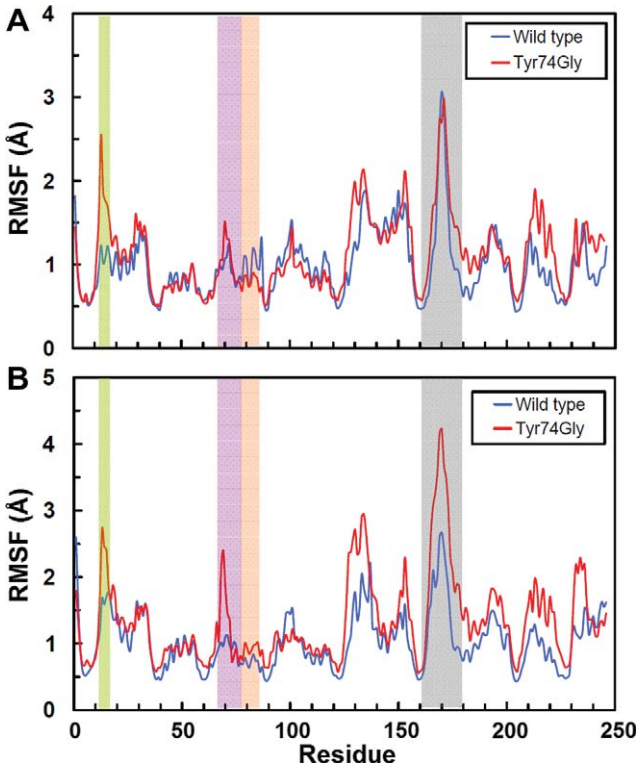

Fig. 3. Root-mean-square fluctuations of Cα atoms about their average positions computed for (A) subunit A and (B) subunit B of wild-type and Tyr74Gly PfTIM dimers. Green, purple and gray shading indicate the location of loops 1, 3 and 6, respectively. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

form. In fact, the structural fluctuations of loop 1 induced by this mutation are comparable to those observed in the isolated monomers. Likewise, Tyr74Gly induces a shift in the backbone fluctuations of loop 3 of subunit B (Fig. 3B). Tyr74Gly also induces an appreciable increase in the dynamics of loop 6 of subunit B.

Perturbation of native contacts at the interface induced by Tyr74Gly

To monitor the effect of Tyr74Gly on intermonomeric contacts between loops 1 and 3, we calculated the fraction of intermolecular native contacts between Cys13 (loop 1) and loop 3 of adjacent subunit (Q_{Cys13}, Fig. 4). We chose Cys13 as (a) in the native state of the dimer, it makes critical contacts with loop 3 and (b) its side chain becomes entirely buried upon dimerization.

At physiological temperature, contacts between Cys13 and loop 3 in the wild-type dimer are held together by favorable electrostatic and van der Waals interactions ($\langle E_{int} \rangle \approx -20$ kcal/mol). Given the inherent flexibility of loop 3, the value of Q_{Cys13} is expected to fluctuate in the trajectory of the wild-type dimer; Figs. 4A and B (blue line) show continuous shifts in time-dependent Q_{Cys13}, in agreement with this assumption. Nevertheless, residue Cys13 of subunits A and B retains at least 50% of its native contacts with loop 3 ($Q_{Cys13} > 0.5$) during the entire 200 ns of simulation. Conversely, when Tyr74 is replaced with glycine, only 52% (subunit A) and 69% (subunit B) of the total simulation time satisfies a $Q_{Cys13} > 0.5$ (Figs. 4A and B, red line).

High-temperature (400 K) 150-ns MD simulations of the dimer were used to accelerate the possible order-to-disorder transitions induced by Tyr74Gly. Despite the use of high temperature, both wild-type and mutant dimers retain most of their secondary and tertiary structure in this timescale. Furthermore, we observed that

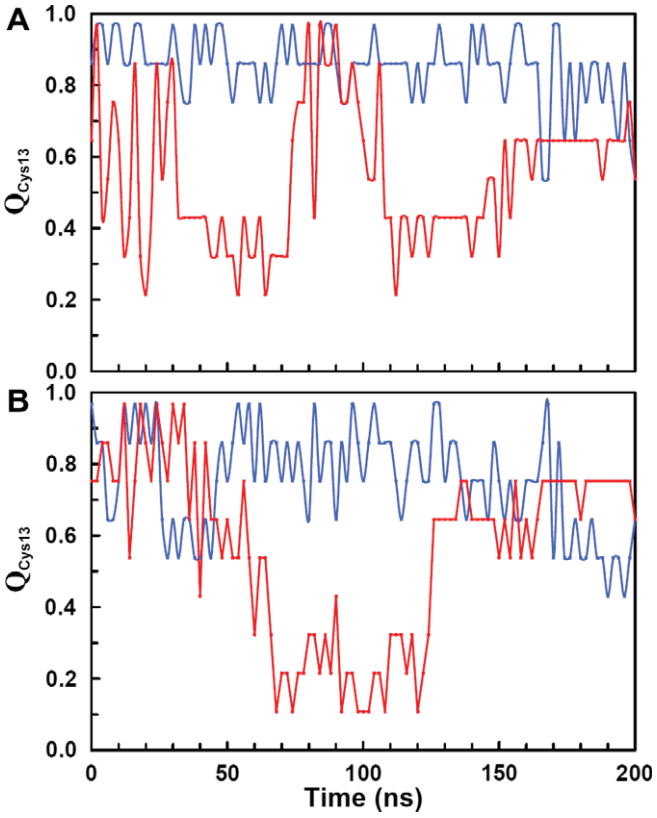

Fig. 4. Fraction of intermolecular native contacts between Cys13 and loop computed for wild-type (blue line) and Tyr74Gly (red line) PfTIM dimers at 310 K. (A) Fraction of intermolecular contacts involving Cys13 of subunit A; (B) fraction of intermolecular contacts involving Cys13 of subunit B. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

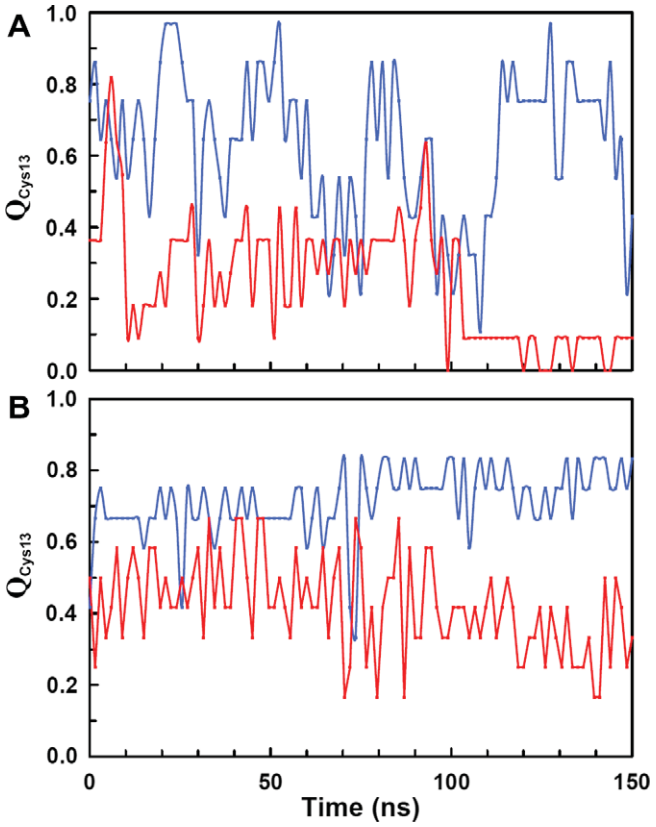

Fig. 5. Fraction of intermolecular native contacts between Cys13 and loop computed for wild-type (blue line) and Tyr74Gly (red line) PfTIM dimers at 400 K. (A) Fraction of intermolecular contacts involving Cys13 of subunit A; (B) fraction of intermolecular contacts involving Cys13 of subunit B. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

high temperature did not produce dramatic changes in the RMSF pattern of most of the protein compared to that obtained at 310 K (data not shown). Nonetheless, the RMSF of loops 3 is 1–1.5 Å larger in the mutant compared to the wild-type, indicating that flexibility of this loop is increased when the side chain of Tyr74 is removed. This increase in flexibility is the result of a large loss of intermolecular contacts between Cys13 and loop 3 (Figs. 5A and B). For instance, only 5% (subunit A) and 20% (subunit B) of the total simulation time at 400 K satisfies a $Q_{Cys13} > 0.5$, whereas native contacts are largely conserved in the wild-type PfTIM (76% and 97% of the total simulation time for subunits A and B, respectively). These observations also correlate with the increase of solvent-accessible surface area (SASA) of Cys13 in the dimeric mutant (Fig. S4, Supplementary material).

Mutation-induced alteration in the backbone conformation of Lys12

The X-ray structure of PfTIM dimer has revealed that active site residue Lys12 lies within a less favorable, albeit generously allowed, region of the Ramachandran plot ($\varphi \approx 50°$, $\psi \approx -150°$) [4]. As similar Ramachandran values have also been observed in other TIM structures, it is assumed that this particular conformation is required for catalytic function. As the structure of the region surrounding Cys13 is affected by the mutant Tyr74Gly, we analyzed the mutation-induced alteration in the backbone conformation of Lys12.

First, we mapped φ and ψ angles of Lys12 onto the Ramachandran plot for each structure in the trajectory of wild-type monomeric PfTIM (Fig. 6). Analysis of the Ramachandran plot revealed

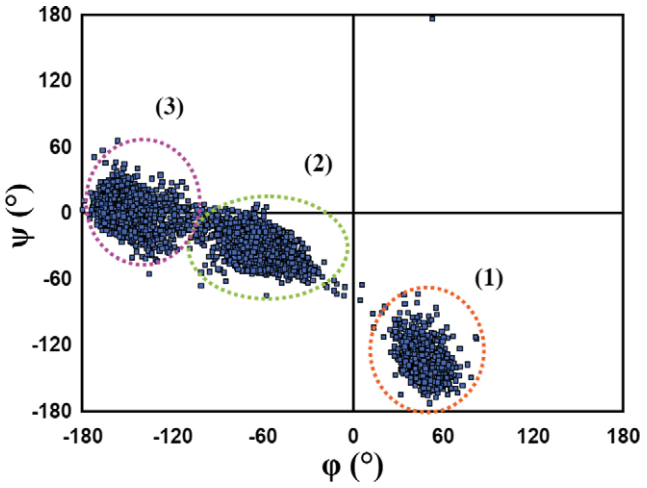

Fig. 6. Ramachandran plot of residue Lys12 in the wild-type monomeric PfTIM. Each oval represents a particular population of backbone conformations in the trajectory (see text for interpretation).

the existence of three populated clusters. The first cluster (Fig. 6, orange[2] oval) corresponds to the Ramachandran values observed

[2] For interpretation of the references in color in this figure, the reader is referred to the web version of this article.

in the crystal structure of the dimer; this conformation remains fairly stable in the trajectory for the first 24 ns. After this period of time, the conformation of Lys12 shifts rapidly to populate a second cluster in the right-handed α-helical region ($\varphi \approx -60°$, $\psi \approx -60°$) and the so-called "bridge region" (Fig. 6, green oval) [26]. Lys12 spends the last 30 ns of simulation populating a backbone conformation having dihedral angles of $\varphi \approx -150°$ and $\psi \approx 0°$ (Fig. 6, purple oval). All these conformational transitions fall within the limits of the energetically allowed regions of the Ramachandran plot. It is worth mentioning that after the initial 24 ns of simulation, Lys12 does not adopt a conformational similar to that in the dimer ($\varphi \approx 50°$, $\psi \approx -150°$). This observation indicates that, in spite of being energetically allowed, this particular backbone conformation does not natively exist in the monomeric form of PfTIM.

Lys12 showed similar conformation patterns in the simulations of wild-type and Tyr74Gly dimers at 310 K. The backbone dihedral angles of Lys12 were $\varphi \approx 50°$ and $\psi \approx -150°$ in each subunit of the wild-type and mutant dimer. Although the free energy of the unbound-to-bound-like conformational transition of Lys12 in the monomer is rather small ($\Delta G \approx 1$ kcal/mol), local stabilizing interactions at the interface of the dimer may increase this free energy difference. As this energy barrier may not be overcome at $T = 310$ K in the sub-microsecond timescale, structural transitions of Lys12 in the dimer were accelerated using MD simulations at $T = 400$ K. Despite the high temperature, Lys12 of each subunit in the wild-type dimer was able to retain its native bound-like conformation (Figs. 7A and B), in agreement with the results obtained for the wild-type dimer at 310 K ($\varphi \approx 50°$, $\psi \approx -150°$). A similar result was obtained for Lys12 of subunit B in the dimeric mutant (Fig. 7D); however, the removal of Tyr74 side chain induces a change in the backbone conformation of Lys12 similar to that observed in the free monomeric wild-type (Fig. 7C).

The effect of disulfide cross-linking at the interface

It has been observed that the mutation of Tyr74 with cysteine followed by chemical cross-linking with Cys13 produces a PfTIM with similar stability to that observed for the wild-type form of the dimer [10]. This observation suggests that Tyr74 somehow participates in the stability of the interface. To address this issue, we performed two MD simulations of dimeric Tyr74Cys mutants: one where Cys13 and Cys74 are connected by a disulfide bond (oxidized PfTIM) and one where Cys13 and Cys74 are not cross-linked (reduced PfTIM).

In terms of structural stability, reduced PfTIM showed an increase in the RMSD of 1 Å compared to oxidized PfTIM, indicating that cross-linking at the interface has a stabilizing effect. Backbone superimposition of oxidized and reduced PfTIM showed that these structural changes in the latter result from the loss of favorable contacts between loops 1 and 3 (Fig. 8). We further calculated the side chain solvent-accessible surface area (SASA) of Cys13 to evaluate to what extent cross-linking enforces solvent inaccessibility (Fig. 9). Contrary to what it was expected, Cys13 of subunit A becomes more accessible to solvent in the oxidized form of PfTIM compared to its reduced form (Fig. 9A). However, this pattern is

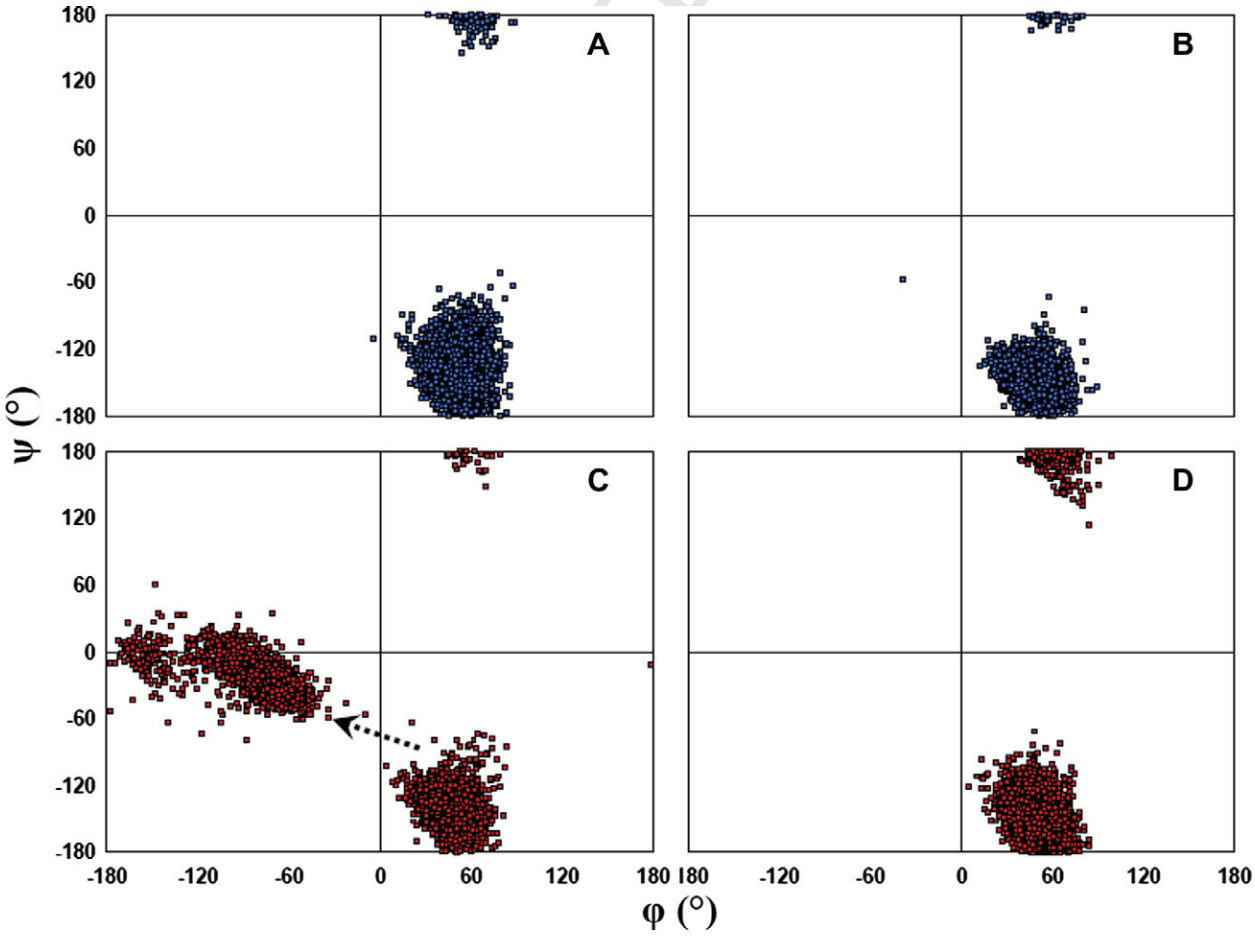

Fig. 7. Ramachandran plot of residue Lys12 in dimeric PfTIM. (A) and (B) correspond to Lys12 of subunits A and B of wild-type PfTIM dimer, respectively. (C) and (D) correspond to Lys12 of subunits A and B of Tyr74Gly PfTIM dimer, respectively.

Please cite this article in press as: L.M. Espinoza-Fonseca et al., Arch. Biochem. Biophys. (2009), doi:10.1016/j.abb.2009.11.009

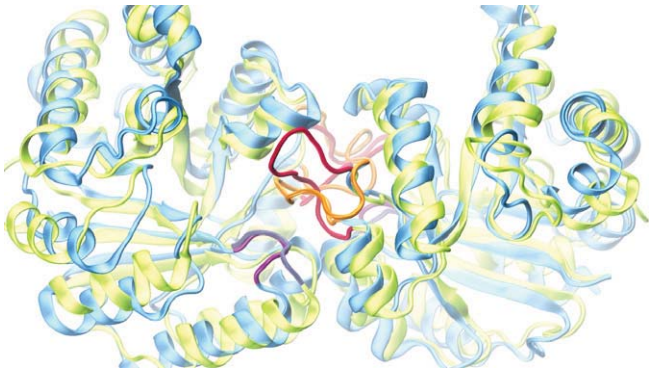

Fig. 8. Superimposition of the final structures (t = 150 ns) of reduced (cyan) and oxidized (green) PfTIM, showing the conformational changes of loops 1 and 3 induced by the absence of cross-linking between mutant Tyr74Cys and Cys13. Loop 1 is shown in blue (reduced) and purple (oxidized), whereas loop 3 is shown in red (reduced) and orange (oxidized). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

different for subunit B, where SASA of Cys13 in the reduced form is larger than that observed in the oxidized PfTIM. The exposure of Cys13 in both oxidized and reduced forms of PfTIM results in the decrease of intermolecular native contacts between loops 1 and 3 (Fig. S5, Supplementary material). Unlike the wild-type form of PfTIM, the oxidized PfTIM is unable to preserve more than 50% for more than a half of total simulation time; comparable results are obtained for reduced PfTIM.

In light of the fact that disulfide cross-linking neither reduces solvent accessibility of Cys13 nor restores the native contacts between loops 1 and 3, it is expected that the structure of loops 1 and 3 is dramatically affected in both oxidized and reduced forms of PfTIM. Superimposition of initial and final structures in the trajectories revealed that, while the native conformation of loop 3 is affected in both cross-linked and non-cross-linked dimers, the structure of loop 1 seems well preserved in the oxidized form. As the backbone conformation of Lys 12 is sensitive to large changes at the interface, we obtained the Ramachandran plots of this residue for each subunit of the oxidized and reduced PfTIM (Fig. 10). To our surprise, the native backbone conformation of Lys12 ($\varphi \approx 50°$, $\psi \approx -150°$) does not change upon disulfide cross-linking (Fig. 10A). In contrast, the absence of a disulfide bond between Cys13 and Cys74 produces a dramatic shift in the conformation of Lys12 (Fig. 10B).

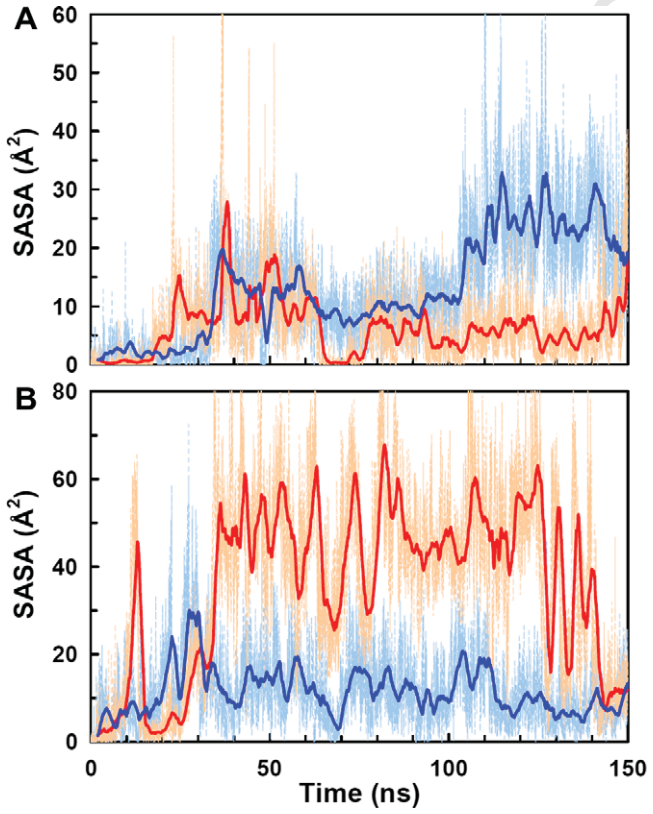

Fig. 9. Solvent-accessible surface area of Cys13 calculated for reduced (red line) and oxidized (blue line) PfTIM. (A) and (B) correspond to Cys13 of subunits A and B, respectively. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

Mobility of Tyr74 in the wild-type monomeric and dimeric PfTIM

We analyzed the evolutionary conservation of Tyr74 using the CONSURF server [27]. We observed that this position is structurally conserved, as its residue variety is only limited to tyrosine and phenylalanine. Considering that recent studies have suggested that the mobility of conserved residues in protein–protein interfaces is restricted [28], we assessed the flexibility of Tyr74 in both monomeric and dimeric PfTIM. Flexibility was evaluated through side chain RMSD following the strategy of Rajamani et al. [29]. To avoid systematic errors caused by translation motions, we calculated side chain RMSD by aligning all heavy backbone atoms of Tyr74; RMSD was calculated for every single structure in the trajectories, and equilibrated structures were used as a reference.

Fig. 11A shows that side chain RMSD of Tyr74 rapidly changes for about 3 Å in the first few nanoseconds of simulation of monomeric PfTIM. Except for the period of time between 40 and 50 ns, the side chain RMSD remains fairly steady throughout most of the simulation time, as expected for a structurally conserved residue. Most of the variations around the average RMSD (Fig. 11A, black line) are the result thermal motions (i.e., the flipping of the aromatic ring in the picosecond timescale). Nonetheless, given the structural symmetry of the side chain, thermal motions have a negligible effect on the average mobility of Tyr74. This observation is supported by structural clustering using a maximum tolerance of 1 Å (Fig. 11B). This clustering procedure showed that Tyr74 side chain visits two conformations during the entire simulation time; however, only one of these two conformations is significantly more populated (Fig. 11B, green sticks).

Tyr74 of PfTIM dimer remains conformational restrained in the subunit A (Fig. 12A). Three conformations are observed for this residue. The most frequent conformation, which resembles to that observed in the crystal structure, is present during the first 125 ns of simulation (~63% of the total simulation time). Although the second (~33% of the time) and third (~4% of the time) conformations of Tyr74 of subunit A deviate from the crystal structure, they appear only as a result of a slight repacking of the aromatic cluster Phe69-Tyr74-Tyr101-Phe102. Tyr74 from subunit B remains very constrained during the entire simulation time (<RMSD> ≈ 0.5 Å), indicating that its packing in the aromatic clusters is not altered in this timescale (Fig. 12B).

Conservation of collective motions in the dimer

The motions of a protein may be broken down into their principal components by principal component analysis (PCA) [30–32]. By combining MD simulations and PCA, Cansu and Doruker showed that dimerization of chicken TIM gives rise to rigid-body motions that are crucial for enzymatic activity [33]. As a large number of critical interactions at the interface involve loops 1 and 3, it is expected that modifications in the geometry of this region will affect

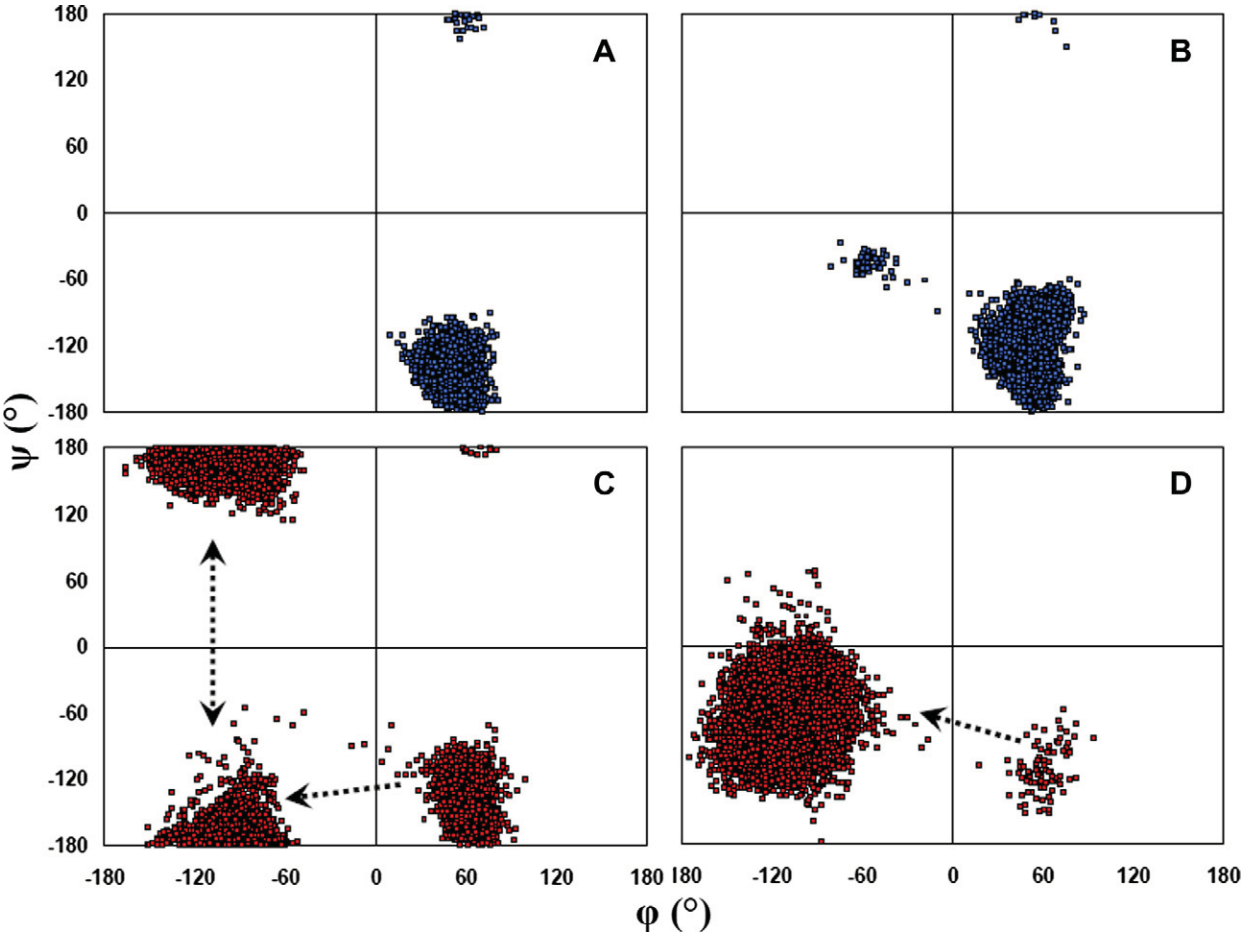

Fig. 10. Ramachandran plot of residue Lys12 in oxidized and reduced PfTIM. (A) and (B) correspond to Lys12 of subunits A and B of oxidized PfTIM dimer, respectively. (C) and (D) correspond to Lys12 of subunits A and B of reduced PfTIM dimer, respectively.

the native rigid-body motions of the dimer. Thus, we extracted the principal components from the wild-type, Tyr74Gly, oxidized and reduced forms of PfTIM.

The first principal component accounts for 37% and 40% of the overall motion in the wild-type and Tyr74Gly PfTIM dimer. For oxidized and reduced PfTIMs, the first principal component describes 44% and 29% of the overall motion of the dimer, respectively. We have plotted the rigid-body motions onto the three-dimensional structure of PfTIM (Fig. 13). The rigid-body motion described by the first (and largest) principal component in wild-type PfTIM consists of a counter-rotation movement of the subunits (Fig. 13A). This motion is consistent with that observed in the wild-type dimeric form of chicken TIM [33]. More importantly, this motion is strongly coupled to the opening/closing of active site loop 6 [33]. In Tyr74Gly dimeric PfTIM, this motion is still present (Fig. 13B); however the homogeneity and magnitude of counter-rotation movement is affected by the removal of Tyr74 side chain. In the reduced form of PfTIM dimer, where the removal of native contacts

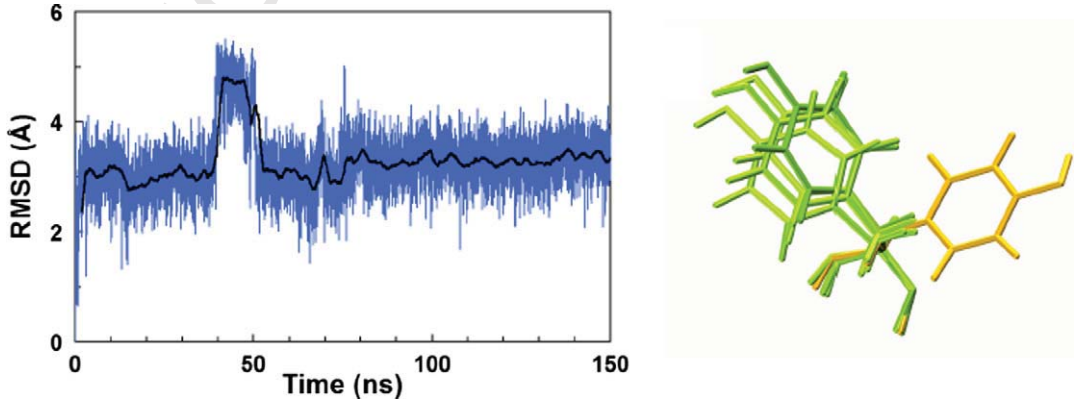

Fig. 11. (A) The root-mean square deviation of Tyr74 of monomeric PfTIM with respect to the structure at $t = 0$ ns. (B) Representative conformations of Tyr74 extracted with the clustering procedure; green structures correspond to the most populated cluster found in the trajectory, whereas the green structure represents the conformation of Tyr74 between 40 and 50 ns. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: L.M. Espinoza-Fonseca et al., Arch. Biochem. Biophys. (2009), doi:10.1016/j.abb.2009.11.009

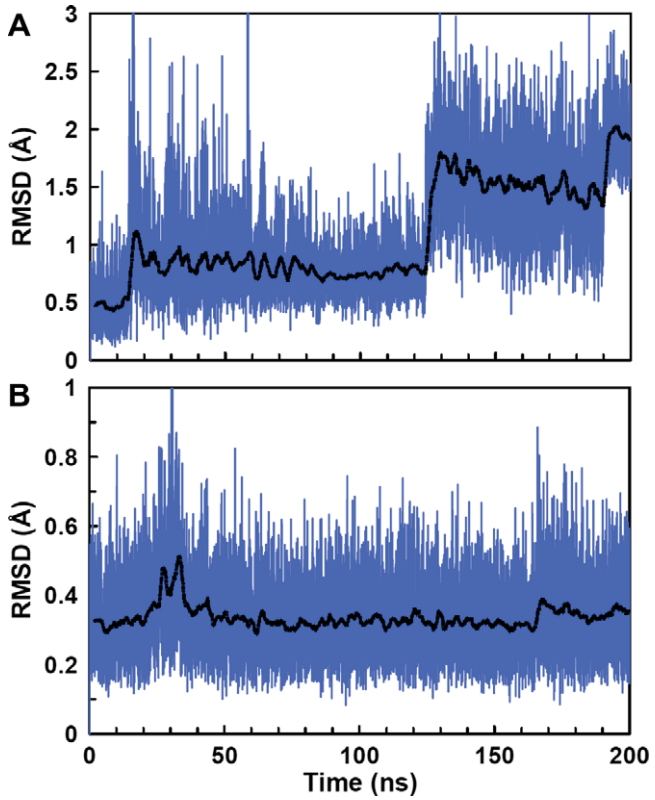

Fig. 12. The root-mean square deviation of Tyr74 of dimeric PfTIM with respect to the structure at $t = 0$ ns. (A) corresponds to subunit A, whereas subunit B is labeled as (B).

and the perturbation of the native conformation of Lys12 are more evident, the native rigid-body motions described by the first principal component are not longer present (Fig. 13D). Surprisingly, the counter-rotation movement observed in wild-type PfTIM dimer appears to be conserved in the oxidized form of PfTIM (Fig. 13C), despite the large loss of native contacts between loops 1 and 3 (Fig. S5, Supplementary material) and increased accessibility of Cys13 (Fig. 9A).

Discussion

Summary of results

In the present study, we have performed MD simulations of PfTIM in the microsecond timescale to unveil the roles of the structurally conserved residue Tyr74 in the formation, stability and function of the PfTIM dimer. Here we studied two mutations of Tyr74 (Tyr74Gly and Tyr74Cys), provided that these mutants have been experimentally analyzed. Simulations of the monomeric PfTIM under physiological conditions revealed that mutant Tyr74Gly does not affect its folding and stability, in agreement with experimental data [11]. We also found that mutation of Tyr74 with glycine does not affect the native local dynamics of loop 3, which is essential for homodimerization. The trajectories of the dimer showed that the mutation does not affect the secondary and tertiary structure of its subunits; however, Tyr74Gly increases the flexibility of loops 1 and 3, decreases the fraction of native contacts between these two loops and alters the native bound-like backbone conformation of active site residue Lys12. Such effects were more noticeable when we used high-temperature MD simulations to accelerate the mutant-induced conformational transitions of loops 1 and 3. Furthermore, the mutant also appears to affect the collective motions of the dimer that are required for catalytic activity. Disulfide cross-linking between Cys13 and the mutant Tyr74Cys appears to restore some of the native structural features of the interface (i.e., the bound-like backbone conformation of Lys12) and the wild-type-like collective motions of PfTIM dimer.

Tyr74 is a ready-made recognition motif for PfTIM homodimerization

In agreement with the experimental data obtained by Maithal et al. [11], we have shown that Tyr74Gly does not affect the folding of monomeric PfTIM. Unfortunately, these experimental studies did not provide any information regarding the local structural changes that this mutation might induce in the monomer. As Tyr74 is located at the tip of loop 3, it is possible that its mutation with glycine may (a) induce a local increase/decrease in flexibility of loop 3 or (b) produce long-range changes in other regions of the monomer via allosteric communication, which may directly affect the conformational entropy of association [34]. However, neither of these scenarios are the case here, as Tyr74Gly does not produce any noticeable shifts in the local dynamics of monomeric PfTIM (Fig. 2). Yet, experiments have shown that monomeric PfTIM holding this mutation is unable to self-associate at low protein concentrations [11]. If Tyr74 does not have an effect in folding and stability of the dimer, how does it influence PfTIM dimerization?

Although a large percentage of monomeric TIM possesses a well-defined three-dimensional structure, approximately 20% of its primary structure remains natively disordered or unstructured. Recent experimental evidence have shown that intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs), which do not possess a well-defined 3D structure under physiological conditions are functional in their native state [35–38]. IDPs/IDRs are often found in protein–protein interactions (PPIs), thus fulfilling a number of key functional roles in the cell. Most importantly, many IDPs/IDRs undergo disorder-to-order transitions upon binding, ranging from the restriction of the dynamics in the primary structure to the formation of a folded structure [39].

Here, we highlight the importance of intrinsic disorder in PfTIM dimerization as two of its regions, loops 1 and 3, are intrinsically disordered in the monomeric form of the enzyme. The intrinsically disordered nature of these loops was assessed by using the FoldUnfold server, which predicts whether a region of a protein in likely to be disordered or not [40,41]. The crystal structure of PfTIM has shown that upon binding, both regions undergo a disorder-to-order transition, where loops 1 and 3 do not become structured but have an important reduction in their accessible conformational states. In our trajectories, we observed that the removal of Tyr74 causes a large increase in the dynamics of loops 1 and 3 and the subsequent loss of native-like contacts between these two IDRs; this result clearly indicates that Tyr74 plays an essential role in the binding-induced disorder-to-order transition of loops 1 and 3. This suggestion is supported by the fact that Tyr74 remains conformationally restricted even in the monomeric (unbound) form of PfTIM (Fig. 11); conformationally restricted residues have been shown to act as anchor residues, which help avoiding kinetically costly structural rearrangements at the interface, allowing for a relatively smooth recognition process [29]. Here, the restricted mobility of the structurally conserved Tyr74 may facilitate the formation of native contacts between loops 1 and 3 without decreasing the binding kinetics via strong native-like contacts [42] and 'fly-casting' mechanism arising from intrinsic disorder [43]. The rigidity of Tyr74 is also essential for specificity in molecular recognition as it provides a ready-made recognition motif [44] for PfTIM homodimerization via the formation of aromatic clusters at the interface. This recognition specificity may be enhanced by other factors such as the type of aromatic–aromatic interaction (i.e., Phe–Phe or Phe–Tyr), the geometrical arrangement of the aromatic

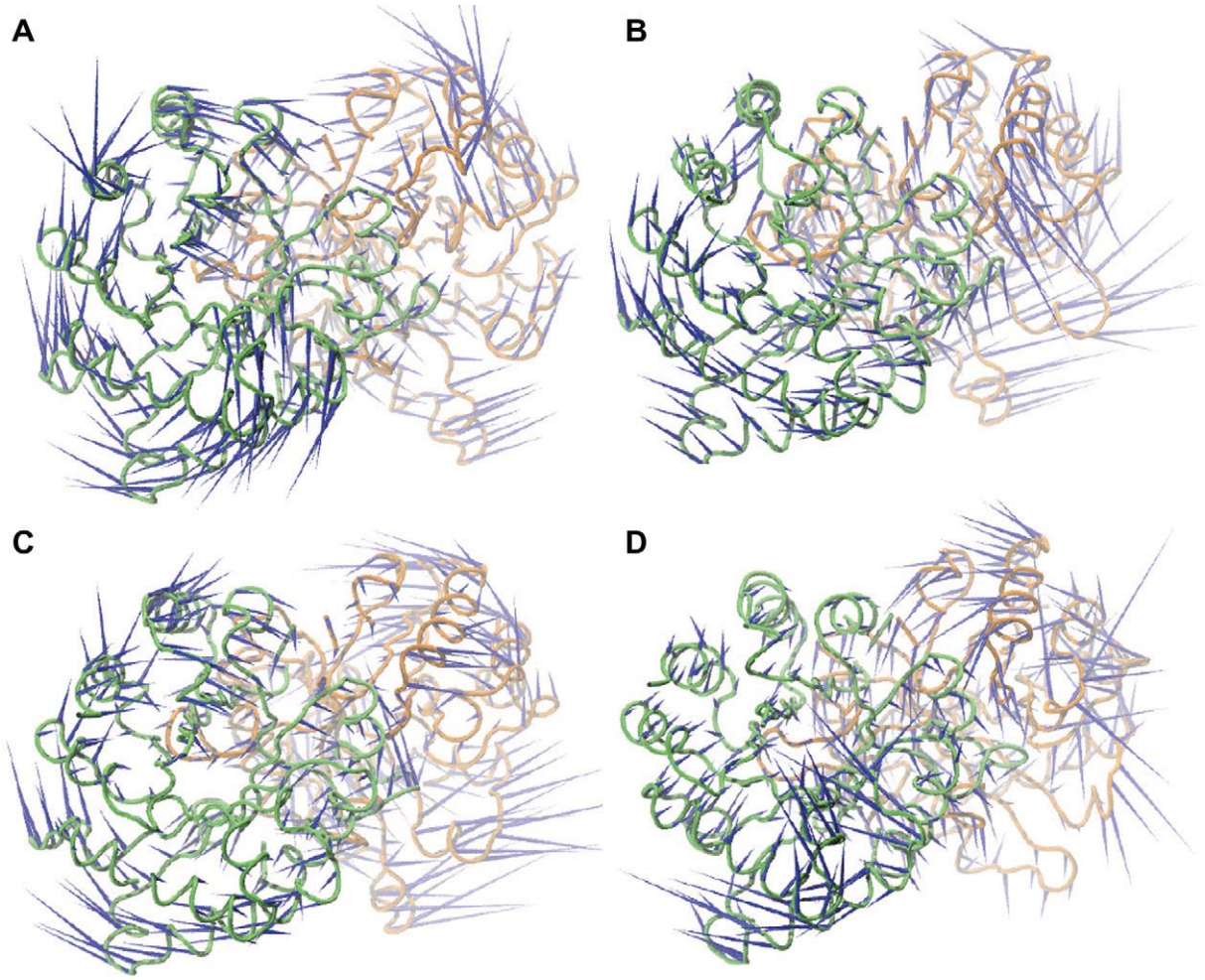

Fig. 13. Collective motions corresponding to the first principal component of (A) wild-type PfTIM, (B) Tyr74Gly PfTIM, (C) oxidized PfTIM and (D) reduced PfTIM. The arrows, starting from the Cα of each residue, show the direction of the movement.

residues (i.e., stacked or T-shaped) and the hydrophobicity of the environment where the interaction takes place [45]. It is important to point out that, according to our simulations, Tyr74 may play a central role in a synergistic binding mechanism between loops 1 and 3. In this mechanism, both conformational selection and induced fit mechanisms participate in the binding of IDPs/IDRs, providing an efficient balance between kinetics and thermodynamics of binding [39]. In this particular case, the side chain of Tyr74 undergoes a conformational selection that produces an induced fit of loop 1, as the bound-like geometry of this loop is not natively populated in the monomer (Fig. 6). This disorder-to-order transition is responsible for the formation of the particular geometrical arrangement of loop 1 that allows Cys13 to fully interact with loop 3 and Lys12 to be locked in a favorable position for catalysis.

The role of Tyr74 in the stability and function of PfTIM dimer

There is strong experimental evidence showing that Cys13 plays a crucial role in the stability of PfTIM dimer. For instance, carboxyl methylation of Cys13 induced the formation of monomeric species, as detected by gel filtration [8]. Furthermore, the mutant Cys13Asp exhibited a reduced stability to denaturants and 7-fold reduction in the enzymatic activity [8]. A similar effect was observed when Cys14 of *Trypanosoma brucei* TIM (equivalent to Cys13 in PfTIM) was either exposed to sulfhydryl reagents or replaced with the other nineteen amino acids. In the case of *T. brucei* TIM, sulfhydryl reagents produced large structural changes and abolition of catalysis [46], whereas mutations induced low stabilities and enzymatic activities compared to the wild-type dimer [47]. Although these studies have acknowledged the importance of Cys13 for the stability of the dimeric form of PfTIM, our simulations and experimental data suggest that Tyr74 plays an equally important role in the stability of PfTIM dimer.

What is the exact role that Tyr74 plays in the stability of the dimer? High-temperature MD simulations revealed that, once the side chain of Tyr74 is removed, Cys13-mediated favorable interactions are loss. This reduction of native contacts is the result of an increase in the accessible surface and a shift in the side chain orientation of Cys13. However, these simulations do not reveal whether this reduction of favorable interactions is the result of a cavity at the interface (as suggested by experiments [10,11]) or a perturbation of the mechanical forces that stabilize Cys13-mediated interactions at the interface. In order to clarify the role of Tyr74 in the stability of the interactions at the interface, we analyzed the trajectories of oxidized and reduced PfTIM dimers. We observed that despite the presence of a cavity at the aromatic clusters of the interface, the reduction of native contacts and the increase of SASA of Cys13, the oxidized (cross-linked) form of PfTIM forces loop 1 to preserve its native-like geometry (Figs. 10A and B). This native bound-like geometry in the dimer is characterized by positive Ramachandran φ values of Lys12 [4], which forces the side chain of Cys13 to interact favorably with loop 3

(Fig. 1B). In the absence of cross-linking, both native contacts and geometry of loop 1 are lost. These results show that disulfide cross-linking mimics the function of Tyr74 at the interface, indicating that Tyr74 possesses a mechanical role via aromatic–aromatic interactions with Phe69, Tyr101 and Phe102.

In addition to this mechanical role, Tyr74 actively participates in the function of PfTIM dimer. Recently, MD simulations of chicken TIM revealed that dimerization gives rise to collective motions that are absent in the monomer [33]. These collective motions, largely described by the first principal component, result in the counter-rotation of the subunits and the closure of loop 6. The results of these MD simulations suggested that these collective motions are necessary for TIM function [33]. Taking into consideration these observations, we extracted the collective motions described by the first principal component of wild-type, mutant, oxidized and reduced PfTIM dimer (Fig. 13). Collective motions described by the first principal component of wild-type PfTIM are the same as for chicken TIM [33], suggesting that such motions are native in the dimeric TIM regardless of species. Surprisingly, Tyr74Gly mutant did not produce large structural changes at physiological temperature, but affected the native collective motions of the dimer (Fig. 13B). What is even more surprising is that the oxidized form of PfTIM is able to maintain the native collective motions observed in the wild-type PfTIM (Fig. 12C). Such motions are no longer present in the reduced form of PfTIM (Fig. 12D). These results suggest that native collective motions that contribute to the catalytic efficiency of PfTIM strongly depend on the preservation of intermolecular interaction between Tyr74 and Cys13.

Hilser and Thompson have shown that site-to-site allosteric coupling is maximized when intrinsic disorder is present in the domains or segments containing one or both of the coupled binding sites [48]. Here, we observed that in the absence of Tyr74, the collective motions coupled to the closure motion of active site loop 6 are importantly affected. Hence, Tyr74 is possibly a link between intrinsic disorder and allosterism in PfTIM, modulating subunit–subunit communication via coupled disorder-to-order transitions of loops 1 and 3, resulting in the collective motions observed in the dimeric form of PfTIM.

Conclusions

By using all-atom MD simulations in the microsecond timescale, we have unveiled the role of Tyr74 in the formation, stability and function of dimeric PfTIM. Our simulations revealed that this structurally conserved residue at the interface is not necessary for the folding and stability of monomeric TIM; however, we observed that the restricted mobility of Tyr74 observed in the dimer is conserved in the monomer, indicating that this residue plays a very important role as an anchor residue for homodimerization. Tyr74 also appears to play an important role in binding by facilitating the disorder-to-order transitions of loops 1 and 3, which allows Cys13 to form favorable interactions with loop 3 and Lys12 to be locked in a favorable position for catalysis. Disulfide cross-linking between mutant Tyr74Cys and Cys13 further revealed that Tyr74 stabilizes the geometry of loop 1 (which contains the catalytic residue Lys12) and the interactions between loops 1 and 3 via aromatic–aromatic interactions with residues Phe69, Tyr101, and Phe102. Principal component analysis revealed that Tyr74 is also necessary to preserve the collective motions in the dimer that contribute to the catalytic efficiency of PfTIM dimer. We conclude that Tyr74 not only plays a role in the stability of the dimer, as previously suggested [10,11], but also participates in the dimerization process and collective motions via coupled disorder-to-order transitions of intrinsically disordered regions, necessary for efficiency in the catalytic function of PfTIM.

Acknowledgments

The authors would like to thank CONACYT and COFAA/SIP-IPN for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2009.11.009.

References

[1] J.M. Herlihy, S.G. Maister, W.J. Albery, J.R. Knowles, Biochemistry 15 (1976) 5601–5607.
[2] W.J. Albery, J.R. Knowles, Biochemistry 15 (1976) 5627–5631.
[3] W.J. Albery, J.R. Knowles, Biochemistry 15 (1976) 5588–5600.
[4] S.S. Velanker, S.S. Ray, R.S. Gokhale, S. Suma, H. Balaram, P. Balaram, M.R. Murthy, Structure 5 (1997) 751–761.
[5] E. Lolis, T. Alber, R.C. Davenport, D. Rose, F.C. Hartman, G.A. Petsko, Biochemistry 29 (1990) 6609–6618.
[6] Z. Zhang, S. Sugio, E.A. Komives, K.D. Liu, J.R. Knowles, G.A. Petsko, D. Ringe, Biochemistry 33 (1994) 2830–2837.
[7] R.K. Wierenga, K.H. Kalk, W.G. Hol, J. Mol. Biol. 198 (1987) 109–121.
[8] K. Maithal, G. Ravindra, H. Balaram, P. Balaram, J. Biol. Chem. 277 (2002) 25106–25114.
[9] S.K. Singh, K. Maithal, H. Balaram, P. Balaram, FEBS Lett. 501 (2001) 19–23.
[10] B. Gopal, S.S. Ray, R.S. Gokhale, H. Balaram, M.R. Murthy, P. Balaram, Biochemistry 38 (1999) 478–486.
[11] K. Maithal, G. Ravindra, G. Nagaraj, S.K. Singh, H. Balaram, P. Balaram, Protein Eng. 15 (2002) 575–584.
[12] M. Banerjee, H. Balaram, P. Balaram, FEBS J. 276 (2009) 4169–4183.
[13] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, J. Comput. Chem. 26 (2005) 1781–1802.
[14] H. Li, A.D. Robertson, J.H. Jensen, Proteins 61 (2005) 704–721.
[15] A.D. MacKerell Jr., D. Bashford, M. Bellott, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586–3616.
[16] A.D. MacKerell Jr., M. Feig, C.L. Brooks 3rd, J. Am. Chem. Soc. 126 (2004) 698–699.
[17] W. Weber, P.H. Hünenberger, J.A. McCammon, J. Phys. Chem. B 104 (2000) 3668–3675.
[18] T. Darden, D. York, L. Pedersen, J. Chem. Phys. 98 (1993) 10089–10092.
[19] U. Essmann, L. Perera, M.L. Berkowitz, J. Chem. Phys. 103 (1995) 8577–8593.
[20] W.F. van Gunsteren, H.J.C. Berendsen, Mol. Phys. 34 (1977) 1311–1327.
[21] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. 14 (1996) 33–38. 27–38.
[22] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, J. Comput. Chem. 26 (2005) 1701–1718.
[23] C.P. Barrett, B.A. Hall, M.E. Noble, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2280–2287.
[24] D.L. Pompliano, A. Peyman, J.R. Knowles, Biochemistry 29 (1990) 3186–3194.
[25] M.E. Noble, J.P. Zeelen, R.K. Wierenga, Proteins 16 (1993) 311–326.
[26] P.A. Karplus, Protein Sci. 5 (1996) 1406–1420.
[27] M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko, N. Ben-Tal, Nucleic Acids Res. 33 (2005) W299–302.
[28] O.N. Yogurtcu, S.B. Erdemli, R. Nussinov, M. Turkay, O. Keskin, Biophys J. 94 (2008) 3475–3485.
[29] D. Rajamani, S. Thiel, S. Vajda, C.J. Camacho, Proc. Natl. Acad. Sci. USA 101 (2004) 11287–11292.
[30] A.E. Garcia, Phys. Rev. Lett. 68 (1992) 2696–2699.
[31] A. Amadei, A.B. Linssen, H.J. Berendsen, Proteins 17 (1993) 412–425.
[32] M.A. Balsera, W. Wriggers, Y. Oono, K. Schulten, J. Phys. Chem. 100 (1996) 2567–2572.
[33] S. Cansu, P. Doruker, Biochemistry 47 (2008) 1358–1368.
[34] K.K. Frederick, M.S. Marlow, K.G. Valentine, A.J. Wand, Nature 448 (2007) 325–329.
[35] P.E. Wright, H.J. Dyson, J. Mol. Biol. 293 (1999) 321–331.
[36] A.K. Dunker, Z. Obradovic, Nat. Biotechnol. 19 (2001) 805–806.
[37] P. Tompa, Trends Biochem. Sci. 27 (2002) 527–533.
[38] V.N. Uversky, Protein Sci. 11 (2002) 739–756.
[39] L.M. Espinoza-Fonseca, Biochem. Biophys. Res. Commun. 382 (2009) 479–482.
[40] O.V. Galzitskaya, S.O. Garbuzynskiy, M.Y. Lobanov, Bioinformatics 22 (2006) 2948–2949.
[41] S.O. Garbuzynskiy, M.Y. Lobanov, O.V. Galzitskaya, Protein Sci. 13 (2004) 2871–2877.
[42] L.M. Espinoza-Fonseca, Biochemistry (2009).
[43] B.A. Shoemaker, J.J. Portman, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 97 (2000) 8868–8873.
[44] S.R. Kimura, R.C. Brower, S. Vajda, C.J. Camacho, Biophys. J. 80 (2001) 635–642.

[45] R. Chelli, F.L. Gervasio, P. Procacci, V. Schettino, J. Am. Chem. Soc. 124 (2002) 6133–6143.
[46] G. Garza-Ramos, N. Cabrera, E. Saavedra-Lira, M. Tuena de Gomez-Puyou, P. Ostoa-Saloma, R. Perez-Montfort, A. Gomez-Puyou, Eur. J. Biochem. 253 (1998) 684–691.
[47] G. Hernandez-Alcantara, G. Garza-Ramos, G.M. Hernandez, A. Gomez-Puyou, R. Perez-Montfort, Biochemistry 41 (2002) 4230–4238.
[48] V.J. Hilser, E.B. Thompson, Proc. Natl. Acad. Sci. USA 104 (2007) 8311–8315.