Molecular dynamics simulations of the free and inhibitor-bound cruzain systems in aqueous solvent: insights on the inhibition mechanism in acidic pH

L.V.B. Hoelz**, V.F. Leal³, C.R. Rodrigues⁴, P.G. Pascutti⁵, M.G. Albuquerque⁶, E.M.F. Muri⁷ and L.R.S. Dias**

¹Laboratório de Química Medicinal, Faculdade de Farmácia, Universidade Federal Fluminense (UFF), Rua Mário Viana 523, Santa Rosa, Niterói, RJ 24241-000, Brazil; ²Laboratório ModMoQSAR, Faculdade de Farmácia, Universidade Federal do Rio de Janeiro (UFRJ), Av. Carlos Chagas Filho 373, CCS, Rio de Janeiro, RJ 21941-599, Brazil; ³Laboratório de Modelagem Molecular, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro (UFRJ), Av. Carlos Chagas Filho 373, CCS, Rio de Janeiro, RJ 21941-902, Brazil; ⁴Laboratório de Modelagem Molecular, Instituto de Química, Universidade Federal do Rio de Janeiro (UFRJ), Av. Athos da Silveira Ramos 149, CT, Rio de Janeiro, RJ 21949-900, Brazil

Communicated by Ramaswamy H. Sarma

(Received 20 June 2015; accepted 22 September 2015)

The major cysteine protease of Trypanosoma cruzi, cruzain (CRZ), has been described as a therapeutic target for Chagas’ disease, which affects millions of people worldwide. Thus, a series of CRZ inhibitors has been studied, including a new competitive inhibitor, Nequimed176 (NEQ176). Nevertheless, the structural and dynamic basis for CRZ inhibition remains unclear. Hoping to contribute to this ever-growing understanding of timescale dynamics in the CRZ inhibition mechanism, we have performed the first study using 100 ns of molecular dynamics (MD) simulations of two CRZ systems in an aqueous solvent under pH 5.5: CRZ in the apo form (ligand free) and CRZ complexed to NEQ176. According to the MD simulations, the enzyme adopts an open conformation in the apo form and a closed conformation in the NEQ176–CRZ complex. We also suggest that this closed conformation is related to the hydrogen-bonding interactions between NEQ176 and CRZ, which occurs through key residues, mainly Gly66, Met68, Asn69, and Leu160. In addition, the cross-correlation analysis shows evidence of the correlated motions among Ala110–Thr59 and Asp60–Pro90 regions seem to be crucial for CRZ activity.

**Keywords:** Chagas’ disease; cruzain; cysteine protease inhibitors; molecular dynamics simulation; Nequimed176

1. Introduction

Chagas’ disease or American trypanosomiasis, whose etiological agent is the protozoan parasite Trypanosoma cruzi (T. cruzi), affects about 6–7 million people worldwide, mostly in Latin America, and continues to be an important global health concern (Arce-Fonseca, Rios-Castro, Carrillo-Sánchez, Martínez-Cruz, & Rodríguez-Morales, 2015; Clayton, 2010; Ferreira et al., 2014; Gazos-Lopes et al., 2014; Rogers et al., 2012; Vital, Arribas, & Trossini, 2015; World Health Organization, 2015). T. cruzi has a complex life cycle, transiting between mammalian hosts and a hematophagous reduvid bug (insect vector), which bites and feeds on mammalian blood. Thus, the insect deposits infected feces onto the mammalian skin, allowing the transmission stage of the parasite, the trypomastigote form, to invade through the bite wound and mucous membranes (Clayton, 2010; Sajid et al., 2011). Once within a cell, the parasite transforms into the non-flagellar replicative amastigote form. Infection of visceral organs causes a number of mega syndromes, notably in the heart, colon, and esophagus, but heart failure is more common in infected patients (Clayton, 2010; Sajid et al., 2011). Before to host cell rupture, the amastigote form transform back into trypomastigote that can either re-invade a new host cell or be taken up via a blood meal by the insect vector (Bhattacharya, Hall, Li, & Vaidehi, 2008; Calderano et al., 2014; Cazzulo, Stoka, & Turk, 2001; Clayton, 2010; Sajid et al., 2011; Sayè, Miranda, di Girolamo, Camara, & Pereira, 2014).

In all life stages of T. cruzi, it has been reported that cysteine proteases are ubiquitously expressed, playing an important role in differentiation, cell invasion, intracellular multiplication, and immune evasion (Cazzulo et al., 2001; Doyle, Zhou, Engel, & McKerrow, 2007; Eakin, Mills, Harth, McKerrow, & Craik, 1992; Engel, Doyle, Hsieh, & McKerrow, 1998; Ferreira et al., 2014).

*Corresponding authors. Email: lucashoelz@yahoo.com.br, lucashoelz@id.uff.br (L.V.B. Hoelz); ldias@vm.uff.br (L.R.S. Dias)

© 2015 Informa UK Limited, trading as Taylor & Francis Group
Cruzipain or cruzain (EC: 3.4.22.51; CRZ) is the major cysteine protease of *T. cruzi*, but some authors name the native enzyme as cruzipain and the recombinant form as cruzain (Duschak & Couto, 2009). The CRZ catalytic domain, which shows a typical architecture common to the mammalian papain-like superfamily, contains 215 amino acids, including the catalytic triad residues, namely Cys25, His162, and Asn182 (CRZ numbering; Figure 1(A) and (B)) (Powers, Asgian, Ekici, & James, 2002; Sajid et al., 2011).

The CRZ genes encode a signal peptide (18 residues), a propeptide (104 residues), and the mature enzyme (345 residues), composed by the catalytic domain and a carboxy-terminal extension (C-T, 130 residues) (Duschak & Couto, 2009). Native CRZ has a mixture of isoforms that prevented determination of its 3D structure by X-ray diffraction, but the 3D structures of the catalytic domain of the recombinant enzyme (without the C-T extension) in complex with diverse inhibitors (peptide and non-peptide; irreversible and reversible) have been determined (Duschak & Couto, 2009).

CRZ is very active in epimastigote form and is found in acidic lysosome-like organelles called reservosomes (Duschak & Couto, 2009). Because of its essential role into the parasite cell, CRZ has been described as a therapeutic target for Chagas’ disease (Ferreira et al., 2014; Rogers et al., 2012; Wiggers et al., 2013).

Therefore, a series of CRZ inhibitors has been studied, including irreversible covalent inhibitors, reversible covalent inhibitors, non-covalent inhibitors, and non-interacting compounds (Lee et al., 2012; McKerrow, Rosenthal, Swenerton, & Doyle, 2008; Rogers et al., 2012; Vital et al., 2015). Recently, Wiggers and co-workers have selected by virtual screening in the ZINC database (Irwin & Shoichet, 2005) a non-peptidic inhibitor, 2-[2-(1H-1,2,4-triazol-5-ylsulfonyl)ethanoylamino]thiophene-3-carboxamide (Nequimed176, NEQ176; ZINC code: 07603518; PubChem code: CID 40701156; PDB ID: 1RV; Figure 1(C)) that non-covalently binds to CRZ (PDB ID: 4KLB) with affinity values in the low micromolar range, presenting trypanocidal activity (Wiggers et al., 2013). Nevertheless, the structural and dynamic basis for CRZ inhibition, promoted by NEQ176 and

---

Figure 1. Cartoon of the cruzain 3D structure (PDB ID: 4KLB) showing the composition of (A) α-helices (H1-H4, purple), β-sheets (B1-B9, yellow), and (B) loops (L1-L10). (C) Close view of the Nequimed176 (NEQ176, PDB ID: 1RV) binding site showing the interactions between inhibitor and enzyme.

Note: The catalytic triad (Cys25, His162, and Asn182, CRZ numbering) and the bound inhibitor NEQ176 (atom color code: carbon, gray; nitrogen, blue; oxygen, red; sulfur, yellow) are shown in each representation of enzyme structure.
others related inhibitors (i.e. non-peptide, non-covalent, and reversible), is still a matter of debate.

Although the majority of cysteine protease inhibitors binds into the active site and blocks the access to their target protease, they do not bind in a strictly substrate-like manner (Farady & Craik, 2010; Laskowski & Kato, 1980). Instead, they interact with the cysteine protease subsites and catalytic residues in a non-catalytically competent manner (Farady & Craik, 2010). This view is based upon the observation that static X-ray cysteine proteases structures, several in free form and complexed to inhibitors (mainly covalent inhibitors), show little differences in their conformation, as measured by the deviation of backbone dihedral angles (Gaspari, Varnai, Szappanos, & Perczel, 2010; Laskowski & Qasim, 2000; Ratia et al., 2008). However, heteronuclear NMR (Lee et al., 2012) and molecular dynamics (MD) simulation (Capriles & Dardenne, 2007; Chaudhuri et al., 2011; Durrant, Keranen, Wilson, & McCammon, 2010) studies have revealed that the irreversible covalent inhibitors in complex to the cysteine protease are flexible on the picoseconds–nanoseconds (ps-ns) timescale (Clarkson & Lee, 2004; Gaspari et al., 2010; Hoelz et al., 2011, 2012; Kramer et al., 2014; Kunze et al., 2014; Meher, Kumar, Sharma, & Bandyopadhyay, 2012; Song & Markley, 2003).

In general, these observations are in line with recent experimental and theoretical results emphasizing the role of internal dynamics in cysteine protease function (Gaspari et al., 2010; Henzler-Wildman et al., 2007), and are in apparent contradiction with the rigid lock-and-key model widely accepted for protease inhibition, including cysteine proteases (Laskowski & Qasim, 2000).

Thus, hoping to contribute to this ever-growing understanding of the CRZ dynamic behavior before and after binding to the inhibitor (i.e. in the free and inhibitor-bound enzyme), we have performed MD simulations using two systems in aqueous solvent under pH 5.5 (acid): the free CRZ (i.e. the apo form, CRZAPO) and CRZ bound to the NEQ176 inhibitor (i.e. the NEQ176–CRZ complex, CRZNEQ). This work represents the first study using MD simulation in a timescale of 100 ns to investigate the conformational change involved in the inhibition mechanism of CRZ, an important therapeutic target for Chagas’ disease (Capriles & Dardenne, 2007; Durrant et al., 2010; Rogers et al., 2012).

2. Materials and methods

2.1. Setup and MD simulation

As the starting structure of the MD study, we used the X-ray crystal structure of the catalytic domain (215 residues) of cruzain (CRZ) in complex with the Nequimed176 (NEQ176) inhibitor (Resolution = 2.62 Å), which is available in the Protein Data Bank under PDB ID 4KLB (Wiggers et al., 2013). The free (CRZAPO) and bound (CRZNEQ) aqueous systems under pH 5.5 (i.e. the pH of the reservoir where the CRZ enzyme is located) were prepared deleting all crystallographic water and ion atoms, and, subsequently, adding hydrogen atoms to the CRZ cysteine crystal structure (PDB ID: 4KLB) (Wiggers et al., 2013), using PDB2PQR server (Dolinsky, Nielsen, McCammon, & Baker, 2004; Rostkowski, Olsson, Sondergaard, & Jensen, 2011). Thus, all the amino acids were modeled in the protonation state they exhibit as free amino acids in water under pH 7.0, with the exception of Asp50 and Glu151, which were modeled in their non-dissociated forms, and all histidine residues, which were considered as protonated.

Additionally, the hydrogen atoms were added to the complexed inhibitor (PDB ID: 1RV; NEQ176) (Wiggers et al., 2013) using Marvin Suite v.5.1.1.5 (ChemAxon Inc.), taking into account the protonation state under pH 5.5. The construction of the ligand topology was made using MKTOP program (Ribeiro, Horta, & de Alencastro, 2008) and the partial charges were calculated at Resp Esp charge Derive (R.E.D.) server (Vanquelef et al., 2011).

The MD simulations were carried out using the GROMACS 4.6.5 package (Hess, Kutzner, van der Spoel, & Lindahl, 2008). The CRZ enzyme was solvated within a cubic box (box length 74.5 Å) containing 12,823 TIP4P water molecules (Jorgensen et al., 1996) and counterions. All simulations were carried out in the NpT ensemble, considering periodic boundary conditions. The temperature was maintained at 300 K, using the Nosé–Hoover thermostat coupling (Nose & Klein, 1983), and isotropic pressure coupling was applied, using the Parrinello–Rahman barostat algorithm (Parrinello & Rahman, 1981). Electrostatic interactions were treated with the Particle-Mesh Ewald method, using a cut-off of 1.0 nm, and Lennard–Jones interactions were also cut off at 1.0 nm (Hess et al., 2008). The LINCS (Hess, Bekker, Berendsen, & Fraaije, 1997) and SETTLE (Miyamoto & Kollman, 1992) algorithms were used to constrain all bonds and water molecules, respectively. The time step for integration was set to 2.0 fs. Initial water relaxation dynamics were carried out during 2.0 ns for system equilibration, keeping the NEQ176–CRZ complex position restrained. Subsequently, production run was performed during 100 ns, without position restraints.
2.2. Analyses of the protein structure and dynamics in the free and bound systems

The protein structure and dynamics in the CRZAPO and CRZNEQ aqueous systems were evaluated by the analyses of the secondary structure (SS) elements, root mean square deviations (RMSD) and fluctuations (RMSF), and radius of gyration (Rg) using the DO_DSSP, G_RMS, G_RMSF, and G_GYRATE programs, respectively, of GROMACS 4.6.5 package (Hess et al., 2008). With the exception of RMSF, calculated for the last 70 ns of simulation time, the other parameters (SS elements, RMSD, and Rg) were calculated for the entire simulation time (100 ns). All the graphs were plotted using XMGRACE 5.1.19 program (Turner, 2005).

2.3. Intermolecular hydrogen bonds analysis in the protein-bound system

The number of distinct hydrogen bonds (H-bond) formed by the NEQ176 inhibitor to the amino acids of the enzyme during the 100 ns of MD simulation of the CRZNEQ system was calculated using the G_HBOND program of GROMACS 4.6.5 package (Hess et al., 2008).

2.4. Analysis of the correlated movements in the free and bound protein systems

The residue-by-residue cross-correlation in the CRZAPO and CRZNEQ systems were calculated using the generalized cross-correlation approach applied to the atomic coordinates of the Ca-atoms (backbone), during the last 70 ns of MD simulation in aqueous solvent. This approach is based on the mutual information method developed by the Grubmüller group using the G_CORRELATION program (Lange & Grubmuller, 2006) available in the GROMACS 3.3.3 package (Lindahl, Hess, & van der Spoel, 2001).

2.5. Principal component analysis (PCA) of the protein motion in the free and bound systems

The calculation of the covariance matrix and its diagonalization to get the associated principal components describing the direction and amplitude of the enzyme motions in the CRZAPO and CRZNEQ systems, during the last 70 ns of MD simulation, was performed using the G_COVAR program of GROMACS 4.6.5 package (Hess et al., 2008).

The projection of the simulated structures was conducted on individual trajectories using the Ca-atoms. The G_ANAEIG program of GROMACS package 4.6.5 (Hess et al., 2008) was used for the projection. To avoid general translation and rotation of the enzyme structure, the trajectories were aligned based on the Ca-atoms. Thus, the porcupine plot was built with the MODVECTOR.py script based on the extreme conformations derived from the first principal component, using the G_ANAEIG module. The images of porcupine projections were generated using PYMOL 1.5.0.3 software (Schrödinger, 2013).

3. Results and discussion

3.1. Structure and dynamics of CRZ under pH 5.5 in the free and bound systems

In order to shed some light on the difference in the dynamic behavior of the enzyme before and after binding to the inhibitor (i.e. free and inhibitor-bound enzyme), we have performed the first study using 100 ns of MD simulations of two CRZ aqueous systems under pH 5.5 (acid): CRZAPO (free CRZ) and CRZNEQ (NEQ176–CRZ). In our MD study, the X-ray crystal structure of the CRZ catalytic domain in complex with the NEQ176 inhibitor (PDB ID: 4KLB) (Wiggers et al., 2013) was used as the starting structure. The CRZ catalytic domain contains 215 amino acids, including the catalytic triad residues (Cys25, His162, and Asn182) (Powers et al., 2002; Sajid et al., 2011).

The analysis of CRZ SS elements for each system (CRZAPO and CRZNEQ) shows there is no significant variation in the stability during the entire simulation time (100 ns), since the inhibitor binding does not change the protein fold (see Supplementary Figures S1 and S2).

By comparing the root mean square deviation values of all CRZ Ca-atoms (Ca-RMSD) relative to the starting structures (Figure 2(A)), we have observed that both systems, CRZAPO (black line) and CRZNEQ (red line), achieve the stability at about 30 ns, presenting average RMSD values of 1.5 and 2.5 Å, respectively, in the last 25 ns. In the initial 20 ns of the simulation time, the RMSD values for both systems are practically identical, probably because the two systems (CRZAPO and CRZNEQ) were assembled from the same initial 3D structure (NEQ176–CRZ, PDB ID: 4KLB).

Then, to analyze the structure stability in each segment of the protein, we have calculated the root mean square fluctuations of all CRZ Ca-atoms (Ca-RMSF) for both systems, CRZAPO (black line) and CRZNEQ (red line), over the last 70 ns of simulation (Figure 2(B)). In general, the Ca-RMSF plot shows that there is no major variation in the structure fluctuation with the exception of the loop3 (L3, region between Cys56 and Leu67) and the loop4 (L4, region between Asp87 and Thr107). Interestingly, the presence of the inhibitor increases the L3 and decreases the L4 fluctuation at approximately 1 Å.
In addition, the analysis of the radius of gyration (Rg) (Figure 2(C)) shows only a slight increase of .25 Å for CRZNEQ system (red line) in comparison with CRZAPO system (black line), during the last 40 ns of simulation. Again, as in the RMSD analysis, Rg values for both systems are practically identical in the first half of the simulation time, probably due to the same reason previously discussed.

3.2. Analysis of hydrogen bonding interactions in CRZNEQ system

The NEQ176 inhibitor moved away from its initial position, approximately 2.0 Å (Figure S3), through 100 ns of simulation. During the last 70 ns of the CRZNEQ simulation, the binding mode of NEQ176 was restricted by hydrogen bonds (H-bonds) mainly with Gly66, Met68, Asn69, and Leu160 residues (Figure 3). Interestingly, among the H-bonds described by Wiggers and co-workers for the NEQ176–CRZ complex, only the interaction between the inhibitor and Gly66 seems to be stable (Figure 1(A)–(C)) (PDB ID: 4KLB) (Wiggers et al., 2013).

In the CRZNEQ system (Figure 3), the N1 atom of NEQ176 establishes an H-bond with the backbone carbonyl group of Gly66 (lifetime = 39.68 ns), which is the most persistent H-bond in this system, while N2 interacts with the backbone carbonyl group of Leu160 (lifetime = 27.60 ns). In addition, the inhibitor N4 and N3 atoms make two H-bonds with the backbone amino group of Met68 (lifetimes = 3.10 and 5.45 ns, respectively), where N4 of NEQ176 also interacts with the backbone amino group of Asn69 (lifetime = 19.60 ns) (Figure 3).

Other H-bonds are also observed in the NEQ176–CRZ complex, but during only a few picoseconds of the simulation time (Supplementary Figure S4). These interactions include all the nitrogen atoms of the inhibitor. Thus, N1 forms H-bonds with the Gly66 backbone amino group (lifetime = 3.49 ns), Gln159 side chain NH group (lifetime = .28 ns), and the backbone carbonyl groups of Ser64 (lifetime = .66 ns), Leu160 (lifetime = 1.14 ns), and Asp 161 (lifetime = 1.74 ns). N2 interacts with the Glu208 carboxylate group (lifetime = .004 ns). N3 makes H-bonds with the NH and...
OH groups of Ser61 (lifetimes = 1.0 and .08 ns, respectively), and with the side chain NH groups of Asn69 (lifetime = .02 ns) and Gln159 (lifetime = .82 ns). N4 establishes interactions with the OH and NH groups of Ser61 (lifetimes = .08 and .65 ns, respectively), the NH group of Gly66 (lifetime = 4.02 ns), and the side chain NH groups of Asn69 and Asn70 (lifetimes = .07 and .002 ns, respectively). Finally, N5 forms H-bonds with the side chain carbonyl groups of Asn69 (lifetime = .002 ns) and Gln159 (lifetime = .038 ns), the side chain NH group of Gln159 (lifetime = .01 ns), the carbonyl group of Leu160 (lifetime = .006 ns), and the carboxylate groups of Glu207 and Glu208 (lifetimes = 1.47 and .35 ns, respectively).

Moreover, the O1 atom of NEQ176 interacts with the backbone amino group of Gly66 (lifetime = .14 ns) and the side chain NH group of Gln159 (lifetime = .04 ns), while the O2 atom forms interactions with the NH group of Met68 (lifetime = .69 ns), the side chain NH group of Gln159 (lifetime = .02 ns), and the carboxylate group of Glu208 (lifetime = .83 ns) (Supplementary Figure S3).

3.3. Cross-correlation map analysis in the free and bound protein systems

To gain further insight into the conformational changes of the enzyme before and after binding to the inhibitor, we have investigated the correlation between the motions of the Ca-atom residues by performing cross-correlation maps analysis for both CRZAPO (Figure 4(A)) and CRZNEQ (Figure 4(B)) systems, where these coefficients provide information about the correlation between the fluctuations of the positions of the Ca-atom residues.

In the CRZAPO system (Figure 4(A)), high values are obtained for correlated motions among the Ca-atoms of regions 3 (Ala110–Asp140; β-strands B3 and B4, and helix H4), 4 (Leu160–Gly189; β-strands B6 and B7), and 5 (Glu190–Gly215; β-strands B8 and B9) of CRZ. Although with lower intensity, the movements of regions 1 (Ala1–Thr59; loops L1 and L2, and helices H1 and H2) and 2 (Asp60–Pro90; loop L3, helix H3 and, β-strand B2) are also correlated to the other regions, including the regions 1 and 4 that contain the catalytic triad. On the other hand, the correlated motion analysis reveals an expressive reduction of correlations in the CRZNEQ system (Figure 4(B)). Particularly, the largest reductions of correlations are observed among the regions 3, 4, and 5 of CRZ.

Overall, we suggest that the binding of the inhibitor NEQ176 is capable of disrupting the correlated motions of the CRZ structure, which probably be associated with the activity of catalytic triad and, therefore, enzymatic function.

3.4. PCA in the free and bound protein systems

To better characterize the nature of collective motions observed in the cross-correlation maps of the CRZAPO (Figure 4(A)) and CRZNEQ (Figure 4(B)) systems, PCA were also carried out on the last 70 ns MD simulation trajectories for both aqueous systems. PCA focuses on these dominant motions, which are represented as porcupine projections based on the ensemble of both CRZAPO (Figure 5(A)) and CRZNEQ (Figure 5(B)) systems onto the plane defined by the top two eigenvectors. These projections show the CRZ backbone movements, where the vectors indicate the direction and extent of the motion.

In both systems, the analysis of porcupine projections point out that the movements between the L4 and CRZ core are opposites and present different amplitudes. In the CRZAPO system (Figure 5(A)), the movements provide an opening of the CRZ structure, exposing its active site (open conformation). However, for CRZNEQ system (Figure 5(B)), the enzyme promotes a closing motion of the active site (closed conformation), where the movements of helix 2 (H2) are restricted. Additionally, these results are in line with the radio of gyration analysis, where CRZ in CRZAPO system shows higher values (see Figure 2(C)).

The movements between the enzyme core and L4, related to the open conformation, may be necessary for the biological function of CRZ (Figure 5(A)).
Figure 4. Maps of the correlated motions among the CRZ Cα-atoms of (A) regions 1 (green), 2 (yellow), 3 (orange), 4 (cyan), and 5 (purple) of the enzyme (see the cartoon of the 3D structure on the top) in the (B) CRZAPO (free enzyme) and (C) CRZNEQ (inhibitor-bound enzyme) systems, during the last 70 ns of simulation time.

Note: The strength of the computed correlation between two respective CRZ Cα-atoms is color-coded (see the color bar on the bottom), where highly correlated motions are in red and poorly correlated motions are in blue.

Figure 5. Porcupine plot showing the main concerted motions of the protein structure (L4, loop 4; H2, helix 2) in the CRZAPO (A) and CRZNEQ (B) systems, during the last 70 ns of molecular dynamics simulation in aqueous solution.
NEQ176 binding, mainly by H-bonds, seems to induce conformational changes in CRZ structure, that are not typically sampled in the CRZapo system, and the protein conformational population shifts toward those that can best accommodate the inhibitor interaction (Figure 5(B)).

4. Conclusions

In this work, we have carried out 100 ns of MD simulation to study the structure and dynamic behavior of cruzain (CRZ) in a complex with a non-covalent inhibitor, Nequimed176 (NEQ176). First, by comparing the conformational changes between the aqueous CRZapo (free enzyme) and CRZNEQ systems during the simulation, we have found that, in the apo system, the protein adopts an open conformation, which may represents an active state of CRZ. However, for the CRZNEQ system, the inhibitor binding changes the protein conformation to a closed conformation, probably representing an inactive state. Additionally, the inhibitor binding at the CRZ through the H-bond interactions with key residues (mainly Gly66, Met68, Asn69, and Leu160) is able to stabilize the closed conformation by blocking the enzymatic active site. Then, cross-correlation analysis show evidence of the correlated motions mainly among the Ala110–Asp140, Leu160–Gly189, and Glu190–Gly215 subdomains, as well as the movements related to Ala1–Thr59 and Asp60–Pro90 regions, seems to be crucial for CRZ activity. Therefore, we believe that these findings will significantly facilitate our understanding of the conformational dynamics in the course of CRZ inhibition, where the molecules that stabilize the closed conformation represent drug candidates to CRZ inhibitor and, consequently, an alternative treatment for Chagas’ disease.

List of abbreviations

CD | Chagas’ disease  
CRZ | cruzain  
CRZapo | aqueous solvent system of cruzain in the apo form (ligand free)  
CRZNEQ | aqueous solvent system of cruzain complexed to Nequimed176 (NEQ176–CRZ)  
H-bond | hydrogen bond  
MD | molecular dynamics  
NEQ176 | Nequimed176  
PCA | principal component analysis  
Rg | radius of gyration  
RMSD | root mean square deviation  
RMSF | root mean square fluctuation  
SS | secondary structure  
T. cruzi | Trypanosoma cruzi

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1038/07391102.2015.1100139.

Acknowledgments

The authors would like to thank the Brazilian agencies CNPq, FAPERJ, and CAPES for financial support.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Arce-Fonseca, M., Rios-Castro, M., Carrillo-Sánchez, S. D. C., Martínez-Cruz, M., & Rodríguez-Morales, O. (2015). Prophylactic and therapeutic DNA vaccines against Chagas disease. Parasites & Vectors, 8, 1–7. doi:10.1186/s13071-015-0738-0

Bhattacharya, S., Hall, S., Li, H., & Vaidhehi, N. (2008). Ligand-stabilized conformational states of human beta(2) adrenergic receptor: Insight into G-protein-coupled receptor activation. Biophysical Journal, 94, 2027–2042. doi:10.1529/biophysj.107.117648

Calderano, S., Godoy, P., Soares, D., Sant’Anna, O. A., Schenkman, S., & Elias, M. C. (2014). ORC1/CDC6 and MCM7 distinct associate with chromatin through Trypanosoma cruzi life cycle. Molecular and Biochemical Parasitology, 193, 110–113. doi:10.1016/j.molbiopara.2014.03.004

Capriles, P. V. S. Z., & Dardenne, L. E. (2007). Molecular dynamics simulations of cruzipains 1 and 2 at different temperatures. In M. -F. S. Marie-France & M. E. M. T. Walte (Eds.), Lecture Notes in Computer Science (Vol. 4643, p. 177). Berlin, Heidelberg: Springer. Angra dos Reis, Brazil. doi:10.1007/978-3-540-73731-5_16

Cazzulo, J. J., Stoka, V., & Turk, V. (2001). The major cysteine protease proteinase of Trypanosoma cruzi: A valid target for chemotherapy of Chagas disease. Current Pharmaceutical Design, 7, 1143–1156. doi:10.2174/1381612013397528

Chaudhuri, R., Tang, S., Zhao, G., Lu, H., Case, D. A., & Dardenne, L. E. (2007). Molecular dynamics simulations of cruzipains 1 and 2 at different temperatures. In M. -F. S. Marie-France & M. E. M. T. Walte (Eds.), Lecture Notes in Computer Science (Vol. 4643, p. 177). Berlin, Heidelberg: Springer. Angra dos Reis, Brazil. doi:10.1007/978-3-540-73731-5_16

Clayton, J. (2010). Chagas disease 101. Nature, 465, S4–S5. doi:10.1038/nature09220

Dolinsky, T. J., Nielsen, J. E., Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Research, 32, W665–W667. doi:10.1093/nar/gkh381

Doyle, P. S., Zhou, Y. M., Engel, J. C., & McKerrow, J. H. (2007). A cysteine protease inhibitor cures Chagas’ disease in an immunodeficient-mouse model of infection. Antimicrobial Agents and Chemotherapy, 51, 3932–3939. doi:10.1128/AAC.00436-07
Molecular dynamics simulations of the free and inhibitor-bound cruzain systems in aqueous solvent

Durrant, J. D., Keranen, H., Wilson, B. A., & McCammon, J. A. (2010). Computational identification of uncharacterized cruzain binding sites. *PloS Neglected Tropical Diseases*, 4, e676. doi:10.1371/journal.pntd.0000676

Duschak, V. G., & Couto, A. S. (2009). Cruzipain, the major cysteine protease of Trypanosoma cruzi: A sulfated glycoprotein antigen as relevant candidate for vaccine development and drug target. *A review. Current Medicinal Chemistry*, 16, 3174–3202. doi:10.2174/09298670978802971

Eakin, A. E., Mills, A. A., Harth, G., McKerrow, J. H., & Craik, C. S. (1992). The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*. *Journal of Biological Chemistry*, 267, 7411–7420. Retrieved from http://www.jbc.org/content/267/11/7411.long

Engel, J. C., Doyle, P. S., Hsieh, I., & McKerrow, J. H. (1998). Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection. *Journal of Experimental Medicine*, 188, 725–734. doi:10.1084/jem.188.4.725

Farady, C. J., & Craik, C. S. (2010). Mechanisms of Macromolecular Protease Inhibitors. *ChemBioChem*, 11, 2341–2346. doi:10.1002/cbic.201000442

Ferreira, R. S., Dessoy, M. A., Pauli, I., Souza, M. L., Krogh, R., Sales, A. I. L., … Andricopulo, A. D. (2014). Synthesis, biological evaluation, and structure-activity relationships of potent noncovalent and nonpeptidic cruzain inhibitors as anti-trypanosoma cruzi agents. *Journal of Medicinal Chemistry*, 57, 2380–2392. doi:10.1021/jm401709b

Gaspari, Z., Varnai, P., Szappanos, B., & Perzelc, A. (2010). Reconciling the lock-and-key and dynamic views of canonical serine protease inhibitor action. *FEBS Letters*, 584, 203–206. doi:10.1016/j.febslet.2009.11.058

Gazos-Lopes, F., Oliveira, M. M., Hoelz, L. V. B., Vieira, D. P., Marques, A. F., Nakayasu, E. S., … Almeida, I. C. (2014). Structural and functional analysis of a platelet-activating lysophosphatidylcholine of *Trypanosoma cruzi*. *PloS Neglected Tropical Diseases*, 8, e3077. doi:10.1371/journal.pntd.0003077

Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M., & Kern, D. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, 450, U913–U927. doi:10.1038/nature06407

Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4, 435–447. doi:10.1021/ct700301q

Irwin, J. J., & Shoichet, B. K. (2005). ZINC—a free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45, 177–182. doi:10.1021/ci049714+

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79, 926–935. doi:10.1063/1.445869

Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *Journal of Physical Chemistry B*, 105, 6474–6487. doi:10.1021/jp003919d

Kramer, M., Halleran, D., Rahman, M., Iqbal, M., Anwar, M. I., Sabet, S., … Yousef, M. (2014). Comparative molecular dynamics simulation of hepatitis C virus NS3/4A protease (genotypes 1b, 3a and 4a) predicts conformational instability of the catalytic triad in drug resistant strains. *PLoS One*, 9, e104425. doi:10.1371/journal.pone.0104425

Kunze, J., Todoroff, N., Schneider, P., Rodrigues, T., Geppert, T., Reisen, F., … Schneider, G. (2014). Targeting dynamic pockets of HIV-1 protease by structure-based computational screening for allosteric inhibitors. *Journal of Chemical Information and Modeling*, 54, 987–991. doi:10.1021/ci400712h

Lange, O. F., & Grubmuller, H. (2006). Generalized correlation for biomolecular dynamics. *Proteins-Structure Function and Bioinformatics*, 62, 1053–1061. doi:10.1002/prot.20784

Laskowski, M., & Kato, I. (1980). Protein Inhibitors of Proteases. *Annual Review of Biochemistry*, 49, 593–626. doi:10.1146/annurev.bi.49.070180.003113

Laskowski, M., & Qamis, M. A. (2000). What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology*, 1477, 324–337. doi:10.1016/S0005-2728(99)00284-8

Lee, G. M., Balouch, E., Goetz, D. H., Lazic, A., McKerrow, J. H., & Craik, C. S. (2012). Mapping inhibitor binding modes on an active cysteine protease via nuclear magnetic resonance spectroscopy. *Biochemistry*, 51, 10087–10098. doi:10.1021/bi301305k

Lindahl, E., Hess, B., & van der Spoel, D. (2001). GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling*, 7, 306–317. doi:10.1007/s008940100045

McKerrow, J. H., Rosenthal, P. J., Swenerton, R., & Doyle, P. (2008). Development of protease inhibitors for protozoan infections. *Current Opinion in Infectious Diseases*, 21, 668–672. doi:10.1097/Qco.0b1013e328315cca9

Meher, B. R., Kumar, M. V., Sharma, S., & Bandypadhyay, P. (2012). Conformational dynamics of HIV-1 protease: A comparative molecular dynamics simulation study with multiple amber force fields. *Journal of Bioinformatics and Computational Biology*, 10, 1250018–1250037. doi:10.1142/S0219720012500187

Miyamoto, S., & Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13, 952–962. doi:10.1002/jcc.540130805

Nose, S., & Klein, M. (1983). Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50, 1055–1076. doi:10.1080/002689783010285
Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics. *Journal of Applied Physics*, 52, 7182–7190. doi:10.1063/1.328693

Powers, J. C., Asgian, J. L., Ekici, O. D., & James, K. E. (2002). Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chemical Reviews*, 102, 4639–4750. doi:10.1021/cr010182v

Ratia, K., Pegan, S., Takayama, J., Sleeman, K., Coughlin, M., Baliji, S., … Mesecar, A. D. (2008). A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 16119–16124. doi:10.1073/pnas.0805240105

Ribeiro, A. A. S. T., Horta, B. A. C., & de Alencastro, R. B. (2008). MKTOP: A program for automatic construction of molecular topologies. *Journal of the Brazilian Chemical Society*, 19, 1433–1435. doi:10.1590/S0103-50532008000700031

Rogers, K. E., Keranen, H., Durrant, J. D., Ratnam, J., Doak, A., Arkin, M. R., & McCammon, J. A. (2012). Novel cruzain inhibitors for the treatment of Chagas’ disease. *Chemical Biology & Drug Design*, 80, 398–405. doi:10.1111/j.1747-0285.2012.01416.x

Rostkowski, M., Olsson, M. H., Sondergaard, C. R., & Jensen, J. H. (2011). Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Structural Biology*, 11, 1–6. doi:10.1186/1472-6807-11-6

Sajid, M., Robertson, S. A., Brinen, L. S., & McKerrow, J. H. (2011). Cruazain: The path from target validation to the clinic. *Advances in Experimental Medicine and Biology*, 712, 100–115. doi:10.1007/978-1-4419-8414-2_7

Sayé, M., Miranda, M. R., di Girolamo, F., Camara, M. D., & Pereira, C. A. (2014). Proline modulates the trypanosoma cruzi resistance to reactive oxygen species and drugs through a novel D,L-proline transporter. *Plos One*, 9, e92028. doi:10.1371/journal.pone.0092028

Song, J. K., & Markley, J. L. (2003). Protein inhibitors of serine proteinases: Role of backbone structure and dynamics in controlling the hydrolysis constant. *Biochemistry*, 42, 5186–5194. doi:10.1021/bi034041u

Turner, P. J. (2005). XMGRACE (Version 5.1.19). Oregon Graduate Institute of Science and Technology. Beaverton, OR: Center For Coastal and Land-Margin Research. http://plasma-gate.weizmann.ac.il/Grace/

Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., … Dupradeau, F. Y. (2011). R.E.D. Server: A web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Research*, 39, W511–W517. doi:10.1093/nar/gkr288

Vital, D. G., Arribas, M., & Trossini, G. H. G. (2015). Molecular modeling and docking application to evaluate cruzain inhibitory activity by chalcones and hydrazides. *Letters in Drug Design & Discovery*, 11, 249–255. doi:10.2174/15701808113106660082

Wiggers, H. J., Rocha, J. R., Fernandes, W. B., Sesti-Costa, R., Carneiro, Z. A., Cheleski, J., … Montanari, C. A. (2013). Non-peptidic cruzain inhibitors with trypanocidal activity discovered by virtual screening and in vitro assay. *Plos Neglected Tropical Diseases*, 7, e2370. doi:10.1371/Journal.Pntd.0002370

World Health Organization. (2015). Fact sheet no 340; Updated March 2015. Retrieved from http://www.who.int/mediacentre/factsheets/fs340/en/

The PyMOL Molecular Graphics System (Version 1.5.0.3). (2013). New York, NY: Schrödinger, LLC. Retrieved from http://www.pymol.org