Acta Cryst D STRUCTURAL BIOLOGY
ISSN 2059-7983

X-ray diffraction and *in vivo* studies reveal the quinary structure of *Trypanosoma cruzi* nucleoside diphosphate kinase 1: a novel helical oligomer structure

Juan Arturo Gomez Barroso,[a,b]* Mariana Reneé Miranda,[c,d] Claudio Alejandro Pereira,[c,d] Richard Charles Garratt[e] and Carlos Fernando Aguilar[a,b]

[a]Laboratorio de Biología Molecular Estructural, Universidad Nacional de San Luis, Ejército de los Andes 950, 5700 San Luis, Argentina, [b]Instituto Multidisciplinario de Investigación en Biología (IMIBIO), Consejo Nacional de Investigaciones Científicas y Técnicas, Ejército de los Andes 950, 5700 San Luis, Argentina, [c]Laboratorio de Parasitología Molecular, Instituto de Investigaciones Médicas (IDIM), Universidad de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina, [d]Instituto de Investigaciones Médicas A. Lanari, Facultad de Medicina, Universidad de Buenos Aires, Buenos Aires, Argentina, and [e]Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlense No. 400, São Carlos, São Paulo 13566-590, Brazil. *Correspondence e-mail: jagomez@unsl.edu.ar

Received 3 August 2021
Accepted 25 October 2021

Edited by G. Cingolani, Thomas Jefferson University, USA

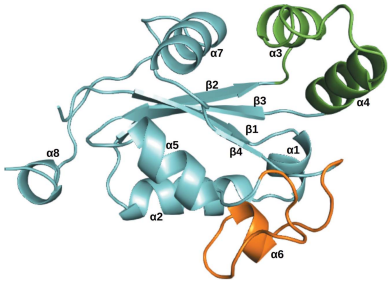
Keywords: quinary multi-hexameric X-ray structure; *Trypanosoma cruzi*; nucleoside diphosphate kinases; *in vivo* protein granules.

PDB reference: *Trypanosoma cruzi* nucleoside diphosphate kinase 1, 7mk0

Supporting information: this article has supporting information at journals.iucr.org/d

Trypanosoma cruzi is a flagellated protozoan parasite that causes Chagas disease, which represents a serious health problem in the Americas. Nucleoside diphosphate kinases (NDPKs) are key enzymes that are implicated in cellular energy management. TcNDPK1 is the canonical isoform in the *T. cruzi* parasite. TcNDPK1 has a cytosolic, perinuclear and nuclear distribution. It is also found in non-membrane-bound filaments adjacent to the nucleus. In the present work, X-ray diffraction and *in vivo* studies of TcNDPK1 are described. The structure reveals a novel, multi-hexameric, left-handed helical oligomer structure. The results of directed mutagenesis studies led to the conclusion that the microscopic TcNDPK1 granules observed *in vivo* in *T. cruzi* parasites are made up by the association of TcNDPK1 oligomers. In the absence of experimental data, analysis of the interactions in the X-ray structure of the TcNDPK1 oligomer suggests the probable assembly and disassembly steps: dimerization, assembly of the hexamer as a trimer of dimers, hexamer association to generate the left-handed helical oligomer structure and finally oligomer association in a parallel manner to form the microscopic TcNDPK1 filaments that are observed *in vivo* in *T. cruzi* parasites. Oligomer disassembly takes place on the binding of substrate in the active site of TcNDPK1, leading to dissociation of the hexamers. This study constitutes the first report of such a protein arrangement, which has never previously been seen for any protein or NDPK. Further studies are needed to determine its physiological role. However, it may suggest a paradigm for protein storage reflecting the complex mechanism of action of TcNDPK1.

1. Introduction



Nucleoside diphosphate kinases (NDPKs; EC 2.7.4.6) are enzymes involved in cellular energy management. They catalyse the transfer of the γ-phosphate from nucleoside triphosphates to nucleoside diphosphates (Parks & Agarwal, 1973), hence maintaining ribonucleotide/deoxyribonucleotide intracellular pools, according to the following reaction,

$$XTP + YDP \longleftrightarrow XDP + YTP.$$

NDPKs can be classified into two groups: one composed of canonical NDPKs and the other of divergent NDPKs. Canonical NDPKs are well conserved enzymes that are present in all organisms. The monomer has a molecular mass of ∼16 kDa

but they are functional as oligomers, for example as hexamers or tetramers depending upon their eukaryotic or prokaryotic origin (Georgescauld *et al.*, 2020). The other group corresponds to large proteins containing one or more NDPK motifs with less characterized N- or C-terminal extensions (Miranda *et al.*, 2011).

Canonical NDPKs possess multiple functions in addition to their well known role in nucleotide homeostasis. In animals, NDPKs are involved in metastasis and invasion of cancer cells (Boissan *et al.*, 2009), probably by acting as a component of various signaling pathways through transcriptional regulation and cytoskeleton interactions (Boissan *et al.*, 2009). They are also involved in cell development and DNA repair (Dorion & Rivoal, 2018; Puts *et al.*, 2018). In plants, NDPKs are involved in carbohydrate metabolism since they participate in the synthesis of cellulose and other cell-wall carbohydrates (Dorion & Rivoal, 2018). In bacteria, they are involved in the fidelity of DNA replication, as deviations in dNTP concentrations produce nucleotide misincorporation in DNA. In addition, they participate in microbe–host interactions, resulting in pathogen adaptation inside the host (Puts *et al.*, 2018; Yu *et al.*, 2016; Sun *et al.*, 2013). Other NDPKs from pathogens such as parasites are involved in proliferation (Castellanos-Gonzalez *et al.*, 2019). They act as virulence factors (Kolli *et al.*, 2008) and participate in responses to DNA damage (Reigada *et al.*, 2020). Despite many studies, the multiple functions of these enzymes are not fully understood.

Trypanosoma cruzi is a flagellated protozoan parasite that causes Chagas disease, an illness that affects millions of people in the Americas (http://www.who.int/entity/mediacentre/factsheets/fs340/en/). *T. cruzi* possesses TcNDPK1, a short canonical NDPK, and three long-isoform variants: TcNDPK2, TcNDPK3 and TcNDPK4 (Camara *et al.*, 2017). Trypanosomatid parasites depend on NDPKs for nucleotide recycling due to their inability to synthesize purines *de novo* (Berriman *et al.*, 2005). Previous reports have shown that TcNDPK1 has nuclease activity *in vitro* (Miranda *et al.*, 2008). More recently, Reigada *et al.* (2020) demonstrated the participation of TcNDPK1 in DNA-damage responses. Like most eukaryotic NDPKs, TcNDPK1 forms hexamers (Ulloa *et al.*, 1995; Souza *et al.*, 2011). Furthermore, Pereira *et al.* (2014) reported that TcNDPK1 forms dynamic cytosolic non-membrane-bound granular structures that localize adjacent to the nucleus.

Oligomers often have advantageous properties over protein monomers, such as increased stability, new binding sites, intersubunit cooperativity and allosteric regulation. Hence, oligomerization can account for functionality. The most common homo-oligomeric states are dimers and tetramers, with less frequent arrangement into supramolecular oligomers. A typical example is the polymerization of dimers of α- and β-tubulin into microtubules.

Quinary structure is the fifth level of protein complexity, in addition to protein primary, secondary, tertiary and quaternary structures. Quinary structure is the result of an evolution in protein structure to allow proteins to navigate the complex crowdedness of the cellular environment. These complexes are stable *in vivo* but are unstable *in vitro*. The protein–protein interactions are 'inherently transient' and imply rapid kinetics in addition to low stability (Wirth & Gruebele, 2013; McConkey, 1982).

Some homomeric quinary associations have recently been reported, such as those of inosine 5-monophosphate dehydrogenase (IMP) and cytosine triphosphate synthase (CTPS), which have been reported as filamentous polymers in cells (Simonet *et al.*, 2020), and monoclonal antibodies (Chen *et al.*, 2016).

In many different metabolic pathways, single or multiple enzymes assemble into cellular ultrastructures. Many of these have punctuated localization and several have filamentous localization patterns, suggesting a high degree of order (Lynch *et al.*, 2020). The formation of higher-order structures is essential for the cellular function of metabolic enzymes that reversibly form filaments and other supramolecular complexes in response to nutrient availability (Lynch *et al.*, 2020). Filamentous enzymes have been identified in most major metabolic pathways, such as nucleotide, fatty-acid and glucose metabolism, and have been described in prokaryotic and eukaryotic organisms. They are structurally diverse and can adopt different conformations, giving rise to diverse three-dimensional filament architectures (Lynch *et al.*, 2020).

Many enzyme filaments form and can also be induced to disassemble in response to signals, including binding to allosteric effectors or ligands. Others are known to be disassembled by the action of buffer components and high ionic strength. Some filamentous enzymes may be capable of responding to environmental conditions without dissociating and reassembling into a new filamentous structure (Park & Horton, 2019).

Filament formation is another layer of enzyme regulation and enables more rapid and more cooperative activation or inactivation of the enzyme. The enzymes thus evolve different biological roles and possible selective advantages. The advantages include the diffusion of substrates and products through the gaps between protomers, allowing substrate selectivity; the localization of the products of the enzymatic reaction to particular locations; the ability to sequester enzymes away from one another and their substrates, modulating metabolic flux; the binding of large substrates; the storage of unneeded enzymes, protecting them against degradation and allowing their rapid redeployment; the protection of active enzymes from degradation or oxidation; the gain of additional functionality; the ability to signal to the cell about the levels of metabolites or other environmental stimuli; the control of water activity; and the buffering of enzymatic activity (Park & Horton, 2019).

The formation of filaments can act as a regulatory mechanism by altering the enzymatic conformation in the polymerized versus nonpolymerized state, with the ability to produce new enzymatic conformations or to generate new regulatory sites at the interfaces of the filaments. Polymerization into filaments may provide a relatively simple, easy-to-evolve mechanism to enhance the regulation of metabolic enzymes. A majority of filamentous metabolic enzymes characterized to date assemble in the active confor-

Table 1
Crystallization.

Method	Hanging-drop vapor diffusion
Plate type	Linbro box
Temperature (K)	277
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	20 m*M* Tris–HCl pH 7.0, 20 m*M* NaCl, 10 m*M* ATP, 10 m*M* $MgCl_2$
Composition of reservoir solution	0.2 *M* $MgCl_2$, 18–20% PEG 3350
Volume and ratio of drop	1.5 µl, 1:1
Volume of reservoir (ml)	1

mation. This suggests that filament formation could increase pathway flux under conditions of high metabolic demand (Lynch *et al.*, 2020).

In the present work, we describe the X-ray diffraction structure of TcNDPK1 at 3.5 Å resolution. Our work revealed the quinary structure of TcNDPK1 as a supramolecular helical oligomer for the first time.

Moreover, enzymatic *in vitro* assays of TcNDPK1 showed the spontaneous dissociation of the filaments in the presence of different nucleotides. These filamentous structures were similar to the structures described for the human counterparts NM23-H1/H2 (Bosnar *et al.*, 2009).

The combination of the structural data with the results of site-directed mutagenesis studies suggest that the microscopic TcNDPK1 granules that are observed *in vivo* in *T. cruzi* parasites are formed by the association of TcNDPK1 oligomers.

2. Methods

2.1. Crystallization, data collection and processing

The crystallization, data collection and data processing of TcNDPK1 have been reported previously (Gomez-Barroso *et al.*, 2010). A short summary of the previous report follows for clarity.

2.1.1. Protein crystallization. Initial crystallization screening was performed using the sitting-drop vapor-diffusion method against The Classics Suite, The Classics II Suite and The PEGs II Suite (Qiagen). The protein solution used was at 10 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.0, 20 m*M* NaCl or with 10 m*M* ATP and 10 m*M* $MgCl_2$. TcNDPK1 crystals were obtained under various screening conditions using ATP and $MgCl_2$. Solution 70 of The Classics Suite (0.2 *M* ammonium sulfate, 0.1 *M* sodium cacodylate pH 6.5, 30% PEG 8000) and solution 47 of The PEGs II Suite (0.2 *M* $MgCl_2$, 0.1 *M* Tris pH 8.5, 30% PEG 4000) were the best conditions. Initial crystals appeared after 24–48 h. Protein crystal optimization was performed by hanging-drop vapor diffusion with a 3 µl drop and a 1:1 ratio of protein to well solution. This step was performed around the initial conditions, varying the buffer pH, the precipitant concentration and the salt composition and concentration, and using temperatures of 293 and 277 K. The best TcNDPK1 protein crystals were obtained using 10 mg ml^{-1} recombinant TcNDPK1 protein in 20 m*M* Tris–HCl pH 7.0, 20 m*M* NaCl, 10 m*M* ATP, 10 m*M* $MgCl_2$ with 20% PEG 3350, 200 m*M* $MgCl_2$ as the reservoir solution at 277 K (Table 1). After 72 h the protein crystals were cryoprotected by transfer to 20%(*v*/*v*) glycerol containing 20% PEG 3350 and 200 m*M* $MgCl_2$.

Table 2
Data collection and processing.

Diffraction source	National Synchrotron Light Laboratory (LNLS)
Wavelength (Å)	1.45
Temperature (K)	100
Space group	*P*3
a, *b*, *c* (Å)	127.84, 127.84, 275.49
α, β, γ (°)	90, 90, 120
Resolution range (Å)	86.38–3.00
Total No. of reflections	103460
No. of unique reflections	58612
Completeness (%)	92.3
Multiplicity	1.8
$\langle I/\sigma(I)\rangle$	2.7
R_{merge}†	0.341
$R_{p.i.m.}$	0.323
R_{meas}	0.471

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th observation and $\langle I(hkl)\rangle$ is the mean intensity of the reflections.

The crystallization of TcNDPK1 was very arduous and protracted. The crystals had a rod-like shape (Supplementary Fig. S1) and were inherently unstable. They grew for 72 h to a size of 500–1000 µm and then start fraying at both ends. They diffracted to low resolution. Our efforts over a long period of time to improve the quality of the crystals and obtain higher resolution data were unsuccessful.

2.1.2. Data collection and processing. A data set was collected to 3.5 Å resolution using synchrotron X-ray radiation at the National Synchrotron Light Laboratory (LNLS), Campinas, Brazil. The data set was reduced and merged using *MOSFLM* (Leslie, 1992) and *SCALA* from *CCP*4 (Winn *et al.*, 2011). The crystals belonged to the trigonal space group *P*3, with unit-cell parameters $a = b = 127.84$, $c = 275.49$ Å (Gómez Barroso *et al.*, 2010; Table 2).

Calculation of the Matthews coefficient suggested that the most suitable value was 2.58 Å^3 Da^{-1} (corresponding to 52% solvent content), indicating 24 molecules per asymmetric unit. All reported structures of eukaryotic NDPKs are hexamers with *D*3 point-group symmetry. This suggested that the content of the asymmetric unit might be four hexamers. A visual check of the data images during data collection did not show obvious twinning or any other abnormality apart from the poor diffraction, which was reflected in a very high $R_{p.i.m.}$ (0.323) and R_{meas} (0.471). Data processing and statistical analysis of the data were performed with *TRUNCATE* (from *CCP*4), which permits the use of the *H*-test (Yeates, 1988) and *L*-test (Padilla & Yeates, 2003). The *L*-test is considered to be the most robust statistical test for twinning in macromolecular crystallography (Thompson, 2017; Campeotto *et al.*, 2018). The *L*-test prediction is also performed internally in *REFMAC*5 (Murshudov *et al.*, 2011) to estimate the twinning fraction during refinement.

The *L*-test gave a value of 0.447, which would suggest a small extent of twinning. The ideal value for an untwinned crystal is 0.5 and that for a perfect twin is 0.375. The cumulative distribution function of *L* indicated that the observed values were similar to the expected curve for untwinned data

and differed significantly from the expected values for twinned data (Supplementary Fig. S2*a*). Finally, the *L*-test proposed a twinning fraction of 0.11.

The *H*-test cumulative distribution for the *P*3 operators $(-h - k, k, -l)$, $(-h, -k, -l)$ and $(h + k, -k, -l)$ showed values of 0.15, 0.17 and 0.16, respectively. They indicated a twinning fraction of 0.17.

Space group *P*3 combined with the presence of noncrystallographic symmetry and large unit-cell dimensions supports tetartohedral twinning. The *Z* plot (Yeates & Yu, 2008), a cumulative intensity statistics distribution, is a test to estimate twin fractions for tetartohedrally twinned data. The *Z* plot for tetartohedral twinned intensities has a sigmoidal shape. Our results showed that the *Z* plot has an exponential shape that would correspond to untwinned intensities (Supplementary Fig. S2*b*). However, in their article *Tetartohedral twinning could happen to you too* Roversi *et al.* (2012) concluded that

> If the extent of twinning is small and/or is obscured by the presence of noncrystallographic symmetry, and especially when the NCS axes coincide with the directions of twinning (the latter introducing deviations from the intensity statistics used to derive the twinning tests), it can also be the case that twinning can only be confirmed at a stage as late as that of refinement of the model.

Therefore, the possibility of tetartohedral twinning was taken into account in refinement with *REFMAC5*.

2.2. Structure solution

2.2.1. Patterson self-rotation function. The self-rotation function comprises peaks generated by the crystallographic operations and noncrystallographic symmetry (NCS) of the molecules in the asymmetric unit. Therefore, it will show the presence of all rotational and screw symmetry axes in the unit cell, the orientation of these axes and the degree of rotation about these axes required to bring a pair of symmetry-related copies into coincidence. The map is plotted in sections of κ. For instance, peaks on the $\kappa = 180^\circ$ section indicate the presence of all twofold rotation and screw axes, peaks on the $\kappa = 120^\circ$ section indicate the presence of all threefolds, and so on. Each κ section can be thought as a spherical projection in which the peaks mark the locations where rotational symmetry axes enter and exit the sphere. All axes intersect at the center of the sphere. The orientations of the axes in the sphere correspond to the orientations of the *n*-fold symmetry axes in the crystal, but the positions of the *n*-fold symmetry axes in the crystal are not specified.

The Patterson self-rotation function was calculated using *MOLREP* (Vagin & Teplyakov, 2010) in the *CCP*4 suite (Winn *et al.*, 2011). The initial integration radii were grouped into two sets. One is based on the size and shape of the unit cell, 64 and 135 Å, corresponding to half of the unit cell along *a*/*b* and *c*. The second set of radii, 42, 90 and 120 Å, take into account the size of one protein hexamer and four linearly related hexamers. The χ angle was varied from 0° to 180° in steps of 5°.

2.2.2. Molecular replacement and refinement. In X-ray crystallography the relative phases of diffracted X-ray beams are lost when the diffraction pattern is recorded. To solve this problem, phases must be obtained by other methods. Several phasing methods can be used. If the unknown crystal structure has sequence identity to a known protein of similar structure (25% or greater) the most used method is molecular replacement (MR). In this method the phases are derived using the atomic coordinates of the structurally similar protein, which is known as the 'search model'. The importance of phases in producing the correct electron density, or structure, is beautifully illustrated in Kevin Cowtan's *Book of Fourier* (http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html). The calculation of an 'electron-density map' using amplitudes derived from the diffraction of a duck and phases derived from the diffraction of a cat results in a cat: the phases carry much more information. The consequence of using an atomic model to calculate crystallographic phases is model bias. This is the main drawback of the molecular-replacement method.

Another problem with molecular replacement is that often the search model does not have sufficient scattering power to generate a solution with a signal-to-noise ratio that is high enough to be identified. This may happen when there are multiple objects in the asymmetric unit or when the search model is a small fraction (a domain or less) of the actual final structure.

Important parameters in MR are the quality of the search model and the quality of the diffracted intensities, which are reflected in the signal-to-noise ratio.

In the case of TcNDPK1, our experience indicated that model phase bias combined with low-quality diffraction intensities and multiple objects in the asymmetric unit would affect the scattering power negatively, generating solutions with low signal-to-noise ratio that are very difficult to identify as a real solution.

The following criteria were adopted for analysis of the MR results.

(i) The point-group symmetry of the solution ought to be compatible with the NCS observed in the self-rotation function.

(ii) Visual observation of molecule packing.

(iii) A suitable electron-density map with no breaks in main-chain density and clear boundaries between molecules in the asymmetric unit.

Several rotation tests were carried out using different programs in order to compare the signal-to-noise ratio of the rotation–translation peaks output for each program. The best results in general were obtained with the *CCP*4 programs *AMoRe* (Navaza, 1994), *Phaser* (McCoy *et al.*, 2007) and *MOLREP*. *MOLREP* was chosen to carry out molecular replacement. The reason for this decision was that the signal-to-noise ratio of the rotation–translation peaks output by this program was the best. This program has a very versatile multi-copy search option incorporated into the program that does not impose any limitation on the oligomeric structure of the protein either in the number of monomers or in their relative location (Scapin, 2013; Vagin & Teplyakov, 2010).

The rotation and translation functions were carried out with the multi-copy search selected, searching for four hexamers (24 monomers) in the unit cell. Rotation assays were

performed using different search models built using *CHAINSAW* (Stein, 2008). The first model was a monomer of *Dictyostelium discoideum* nucleoside diphosphate kinase (1.8 Å resolution; PDB entry 1npk; 56.2% sequence identity; Moréra *et al.*, 1994). The second and third models used were the trimer and the hexamer, respectively, of the structure of human nucleoside diphosphate kinase b NM23-H2 (PDB entry 3bbc; 64% sequence identity; a hexamer with D3 symmetry solved at 1.7 Å resolution; Dexheimer *et al.*, 2009).

The solution coordinates were refined using *REFMAC*5. The refinement method used followed the suggestions of Isupov & Levedev (2008) on how to refine poor-quality low-resolution data. It consisted of 50 cycles of rigid-body refinement followed by 20 subsequent cycles of restrained refinement and finally a further 50 cycles of rigid-body refinement. Before each refinement step the structure was rebuilt using *Coot* (Emsley *et al.*, 2010). The twin refinement option was used for the rigid-body steps and the NCS restraint option was used for the restrained refinement cycles.

Validation was performed with *Coot*, *RAMPAGE* (Lovell *et al.*, 2003), *PROCHECK* (Laskowski *et al.*, 1993) and *SFCHECK* (Vaguine *et al.*, 1999). *PISA* (Mahdavi *et al.*, 2013), *PIC* (Tina *et al.*, 2007) and *CONTACT* (*CCP*4) were used to analyze inter-protein interfaces. Electrostatic and hydrophobic surfaces were calculated using *PyMOL* (DeLano, 2002) and *UCSF Chimera* (Pettersen *et al.*, 2004). Structural analysis and high-quality images of the protein oligomer were obtained using *PyMOL* and *UCSF Chimera*.

2.3. Parasite cultures

Epimastigotes of the Y strain (discrete typing unit II) were cultured at 28°C in plastic flasks (25 cm^2) containing 5 ml BHT medium (started with 10^6 cells per millilitre) supplemented with 10% fetal calf serum, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (Camargo, 1964). Transgenic parasites were obtained by the electroporation of 3×10^8 cells with 50 µg plasmidic DNA in PBS buffer with 0.1 m*M* $CaCl_2$ and 0.5 m*M* $MgCl_2$ in a Bio-Rad Gene Pulser Xcell (400 V, 500 µF). Selection was made with 500 µg ml^{-1} G418 and the parasites were then maintained with the drug at 200 µg ml^{-1}.

2.4. Cloning

All of the GFP fusions (Table 3) were cloned in the pTREX-L vector described previously (Bouvier *et al.*, 2013). Mutations in the TcNDPK1 gene (TritrypDB TcCLB.508707.200) were obtained by Gene Splicing by Overlapping Extension PCR from the pT-L-N1::G vector (Pereira *et al.*, 2014). The ending primers were N1FF, 5′-TCTAGATGACCAGTGAGCGTACCTTC-3′, and N1FR, 5′-GTCGACTGCAGACTCGTAGACCTGCTT-3′. The overlapping primers were K57RF, 5′-TTTGGCCTCGAGGCCGTTCTAC-3′; K57RR, 5′-GTAGAACGGCCTCGAGGCCAAA-3′; K57AF, 5′-TTTGGCCTCGGCGCCGTTCTAC-3′; K57AR, 5′-GTAGAACGGCGCCGAGGCCAAA-3′; delF, 5′-CAGGCTCAGCAGCACGTTGCGTACTTTTCC-3′; delR, 5′-GGAAAAGTACGCAACGTGCTGCTGAGCCTG-3′; alaF, 5′-GCTGCCGCAGCGGCAGCTGCAGCCGCGGCTGCTGCAGCCGTTGCGTACTTTTCCTCC-3′; alaR, 5′-GCTGCCGCAGCGGCAGCTGCAGCCGCGGCTGCTGCAGCCGTTGCGTACTTTTCCTCC-3′. The PCR products were cloned in the pGEM-T Easy vector (Promega), sequenced and subcloned in the pTREX-L-GDH::G vector (Pereira *et al.*, 2014) in the XbaI and SalI recognition sites (underlined).

Table 3
The TcNDPK1 mutations evaluated for granule formation *in vivo*.

Mutation	Mutated amino acid(s)	Replaced by	Region
N1-K57A	Lys57	Alanine	Helix assembly
N1-K57R	Lys57	Arginine	Helix assembly
N1-del	^{51}YIDELASKPFYKDL^{63}	Deleted	Inter-hexameric loop association
N1-ala	^{51}YIDELASKPFYKDL^{63}	Alanines	Inter-hexameric loop association

2.5. Fluorescence microscopy

Parasites were washed twice with PBS and settled for 20 min onto poly-L-lysine-coated cover slips. They were then fixed at room temperature for 20 min with 4% formaldehyde in PBS, washed three times with PBS and mounted using Vectashield with DAPI (Vector Laboratories). Cells were observed for GFP fluorescence in an Olympus BX60 microscope and the images were recorded with an Olympus XM10 camera. For statistical analysis, 53 random fields (approximately 1000 parasites) were analyzed for the presence of TcNDPK1 granules and soluble enzyme using the *IBM SPSS Statistics* 20 two-way ANOVA and *post hoc* multiple pairwise comparisons using Tukey's test. Photographs of the fields were obtained under the same conditions for all populations (exposure time and filters) and were then analyzed.

2.6. Western blots

The parasites were collected, counted in a hemocytometer chamber, washed with PBS and suspended directly in 1× cracking buffer. A volume corresponding to 5×10^6 parasites was loaded onto 12% SDS–acrylamide gels and transferred onto a PVDF membrane. The membranes were blocked for 1 h in 5% nonfat milk in T-PBS and incubated with rabbit anti-GFP (Molecular Probes) in blocking buffer at 1:5000 dilution. After three washes, the membranes were incubated with peroxidase-conjugated anti-rabbit antibodies diluted 1:5000. The proteins were revealed with Super Signal West Pico Chemiluminescent substrate (Pierce). For native electrophoresis, parasites in 50 m*M* Tris–HCl buffer pH 6.8 were subjected to four cycles of freezing and thawing and were centrifuged. The supernatant was mixed with 5× loading buffer. A volume corresponding to 5×10^6 parasites was loaded onto 8% acrylamide gels without SDS and transferred onto PVDF membranes.

3. Results

3.1. Structure solution and refinement

3.1.1. Patterson self-rotation function. The $\kappa = 180°$ section shows peaks corresponding to twofold axes related by crystallographic and noncrystallographic symmetry as described in

Section 2. When the integration radius used was 137.75 Å (half of the unit cell) the $\kappa = 180°$ section showed two sets of peaks: three intense peaks 60° apart and a set of three less intense peaks 60° apart rotated 30° from the strong peaks. The Matthews coefficient indicated the presence of 24 monomers in the asymmetric unit, equivalent to four hexamers. Each hexamer, which has *D*3 symmetry, would generate three coplanar twofolds at $\kappa = 180°$. These would have a 60° angle between them, which would correspond to one set of peaks observed every 60°. The extra three coplanar twofold axes that appear in our self-rotation could not be explained at this stage of structure solution (Fig. 1). One of the hypotheses analyzed at the time was that the extra set of peaks could be due to crystal twinning. However, it was clear at first sight of the structure solution that the extra set of peaks observed in the self-Patterson map was produced by a 30° rotation of the four hexamers in the asymmetric unit to form the helix (Fig. 5).

3.1.2. Molecular replacement. The use of the monomer and the trimer as a search model did not give acceptable results. The signal-to-noise ratio of the top rotation-function peaks obtained in the runs with these models was not good enough to give a precise orientation. Consequently, the translation function had very low contrast values (<2.5) and did not show any feasible solution. It was then decided to use the hexamer as a search model to increase the scattering power. Finally, after several runs, a rotation and translation search using *MOLREP* and the multi-copy search option, with a radius of integration of 90 Å, gave a solution compatible with the criteria described previously. The rotation function displayed 12 strong equivalent solutions (peak height >9σ) which were related to one another by noncrystallographic symmetry. The top ten peaks were chosen to calculate the translation function (contrast value = 4.54; Supplementary Fig. S3).

3.1.3. Refinement. The solution coordinates were then refined as rigid bodies (resolution 10–3.5 Å) using *REFMAC*5. The *R* and R_{free} final values decreased from 0.493 and 0.489 to 0.268 and 0.277, respectively. It was clear on first sight of the structure solution that the extra set of peaks observed in the self-Patterson map was produced by the 30° rotation of the four hexamers in the asymmetric unit to form the helix (Fig. 5).

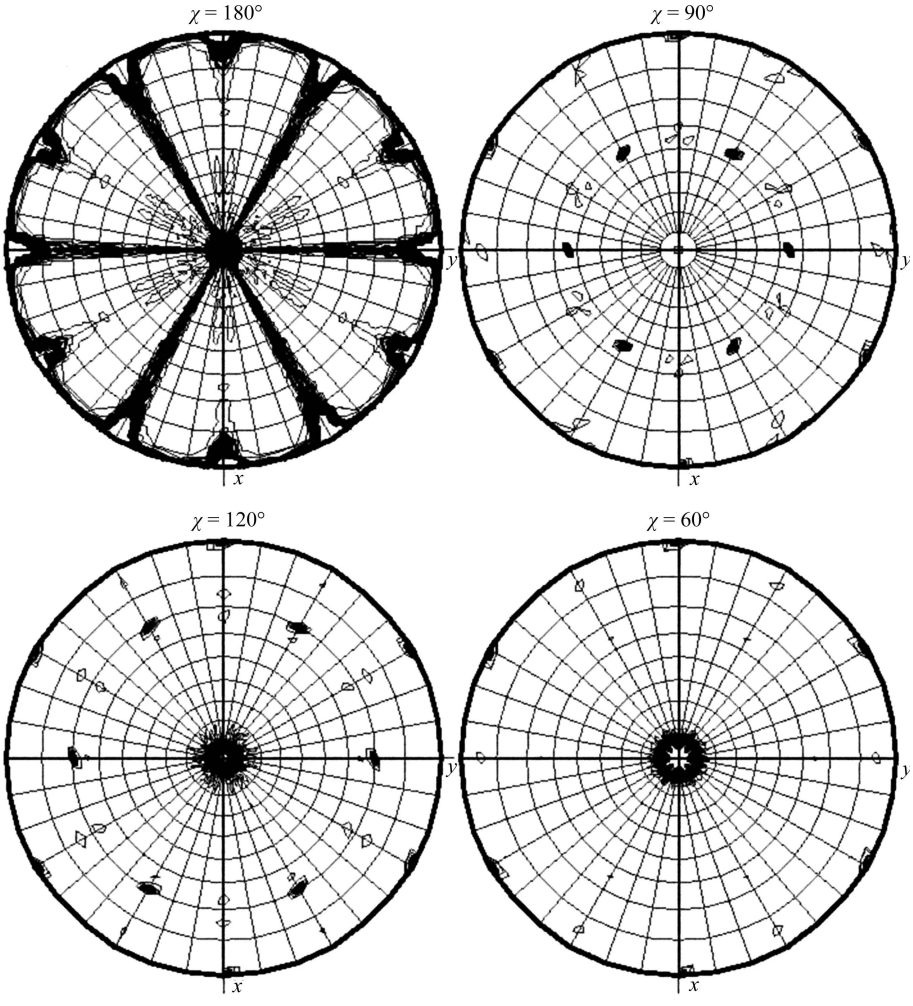


Figure 1
Patterson self-rotation function calculated for the *P*3 native data set using *MOLREP* (*CCP*4): $\kappa = 180°$, $\kappa = 120°$, $\kappa = 90°$ and $\kappa = 60°$. The self-rotation function describes the crystallographic and noncrystallographic symmetry (NCS) of the contents of the asymmetric unit. The radius of integration used was 27.6 Å and the resolution range was 37–5 Å. The $\kappa = 180°$ section showed the crystallographic threefold axis perpendicular to the plane of the figure and two sets of three twofold axes in the plane of the figure. These two sets, of different intensities, are rotated 30° with respect to each other. One set describes the *D*3 point-group noncrystallographic symmetry of the 24 hexamers in the asymmetric unit. The other set describes the 30° rotation of the hexamers with respect to each other in the asymmetric unit, generating the oligomer helix along the *z* axis.

3.2. Monomer structure

The monomeric structure of TcNDPK1 consisted of a core of four β-strands arranged in a $\beta2\beta3\beta1\beta4$ antiparallel sheet topology surrounded by nine α-helices, as observed in canonical NDPK structures (Fig. 2; Georgescauld *et al.*, 2020). The monomer has been classified as belonging to the NDPK superfamily according to *CATH* and *SCOP*. It has an α–β plait topology (code 3.30.70) as classified by *CATH* (Orengo *et al.*, 1997) and a ferredoxin-like fold (code 54861) as classified by *SCOP* (Murzin *et al.*, 1995).

As observed in other NDPKs, the active site of TcNDPK1 comprises a nucleotide-binding site and the nucleophilic histidine. The site accepts two types of substrates, a donor and an acceptor of a phosphate group, depending on its conformational state (Georgescauld *et al.*, 2020). Together

with the Head region, the Kpn loop forms a cleft that harbors the highly positively charged active site that is required for the recognition and binding of negatively charged substrates. 15 residues constitute the active-site cleft, including those involved in nucleotide binding (Lys11, Tyr51, Leu54, Phe59, Leu63, Tyr66, Arg87, Thr93, Arg104, Val111, Gly112, Asn114, Gly118 and Asp120) and the His117 residue needed for nucleotide attack. Other conserved residues are Tyr51, which plays a part in the catalytic mechanism, and Phe59, which is important in making a stacking interaction with nucleotide substrates (Fig. 2). Other important residues, such as Pro95, which is part of the Kpn loop involved in stability, are also conserved (Vieira *et al.*, 2015; Georgescauld *et al.*, 2020).

3.3. Hexamer structure

The formation of a hexamer can be thought of either as a dimer of trimers or a trimer of dimers. It is generally accepted that the strength of the interactions is related to its stability. Furthermore, it is also probable that the most stable association is formed first. Therefore, analysis of these interactions may give us an idea of the dynamics of hexamer formation and the oligomerization steps. Also, it allows characterization of the associations as transient or permanent (Levy & Teichmann, 2013; Dey *et al.*, 2010; Nooren & Thornton, 2003).

Analysis of the protein–protein interactions in the isologous dimer (Monod *et al.*, 1965) showed that the buried interface area has a value greater than 1200 Å. This value is almost double that for the buried interface area generated in the assembly of the heterologous trimer (Monod *et al.*, 1965). The interface of the dimer is formed by the generation of a β-sheet between the monomers, an α-helix–α-helix interaction and the participation of C-terminal amino acids belonging to the Trp141–His144 region (Fig. 3). The formation of the dimer is probably the first step in oligomer assembly (Georgescauld *et al.*, 2020).

Therefore, according to the previous discussion, the second step in the oligomerization would be the formation of the hexamer as a trimer of dimers ordered by *D*3 symmetry. The heterologous association between dimers implies different interactions. Hydrophobic associations were observed between Pro100 and Val88, Val109 and Tyr32, and Val149 and Tyr150. The main ionic interactions were the ion bridges Asp106–Lys29, Asp110–Lys80 and Arg113–Glu151. There is a cation–π bond between Phe107 and Lys29 (Fig. 4).

3.4. Multi-hexameric structure

The third structural level is the formation of the oligomer by the stacking of hexamers parallel to the *z* axis. The hexamers stack together forming a left-handed helix with a 60° rotation

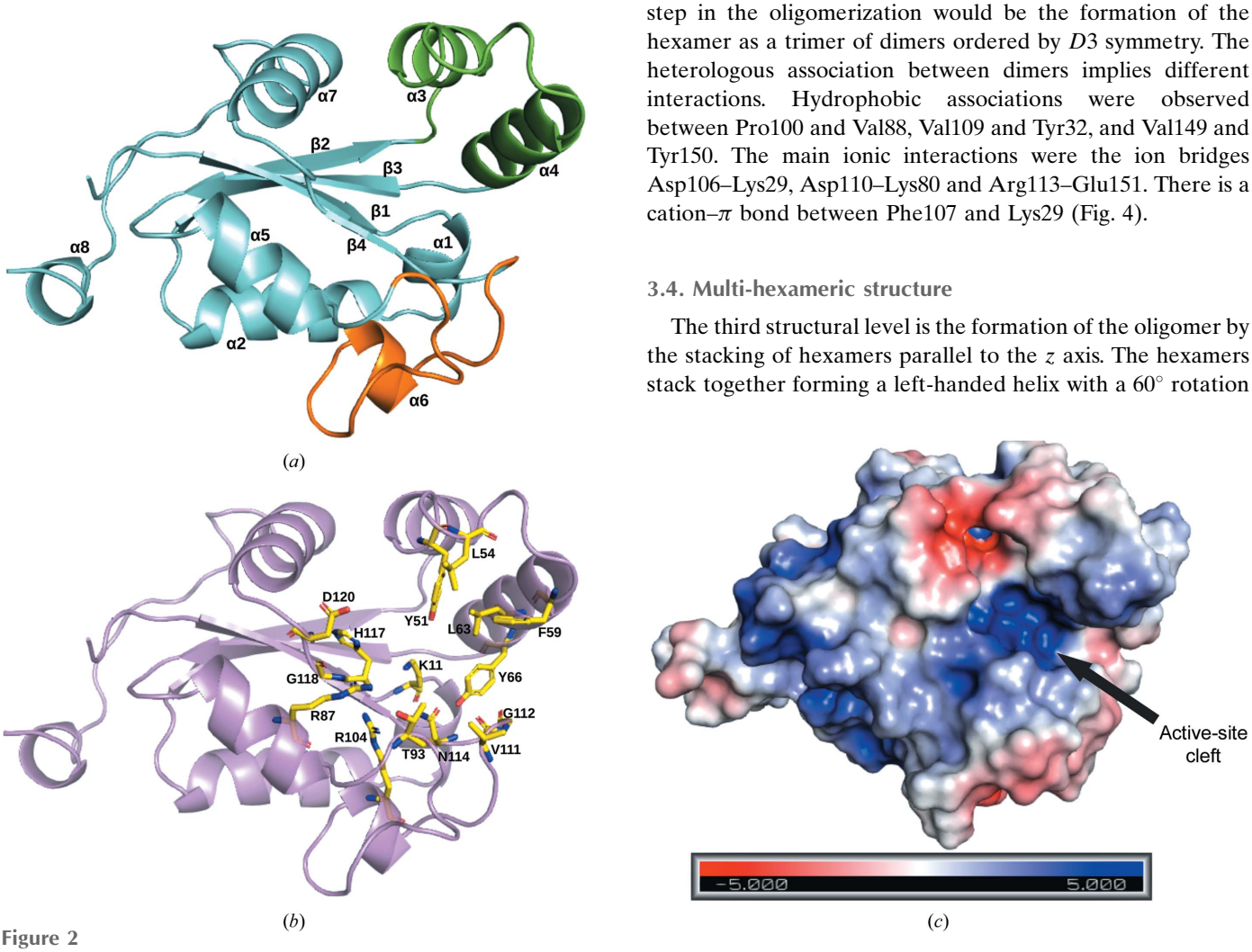


Figure 2
Cartoon representation of the α–β plait topology of the TcNDPK1 monomer. (*a*) α-Helices (α1–α8) and β-sheets (β1–β4) are labeled. Important regions are highlighted: Kpn loop, orange; Head, green. The Kpn loop and Head together harbor the highly positively charged active site required for the recognition and binding of negatively charged substrates. (*b*) Amino acids involved in catalytic activity are shown as sticks and labeled. The most important residues are His117, which is needed for nucleotide attack, Tyr51, which plays a part in the catalytic mechanism, and Phe59, which is necessary for stacking interactions with nucleotide substrates. (*c*) Connolly surface and electrostatic potential calculated using *APBS* and *PyMOL*. An arrow points towards the active-site cleft.

(Fig. 5*a*). The association between the hexamers is mediated by the interaction of three protruding loops on each face of the hexamer ([52]IDLASKPF[59] belonging to the so-called head region in the monomer; Fig. 2).

The loop connects α-helices H3 and H4, forming a helix–loop–helix motif as described previously. The motif is stabilized by a hydrophobic core formed by Ala44, Ala47, Tyr51, Leu54, Ala55, Tyr60, Leu63, Val64 and Phe67. The loops interact though a salt bridge between Lys57 and Asp53 (Fig. 6) plus a hydrogen bond between the Ser56 residues. The bonding surface between the hexamers (595 Å^2) is greatly reduced compared with the binding surfaces associated with dimer and hexamer formation. This value is associated with transitory protein–protein association (Nooren & Thornton, 2003), suggesting that the hexamers bind and unbind dynamically to form the oligomer. An analysis of the electrostatic inter-hexameric surface shows a pattern of complementary charged zones produced by the 60° rotation of the hexamers in the helix, favoring proper orientation of the hexamers.

3.5. Multi-helix oligomerization

A fourth level of organization is observed in which several oligomeric helixes are crystallographically related by the trigonal symmetry axes parallel to *c* in the crystal and each helix is surrounded by six parallel helices (Fig. 5*d*).

3.6. The helix is related to granules observed *in vivo*

As mentioned in Section 1, Pereira *et al.* (2014) reported that TcNDPK1 forms large granules in the parasites. Different mutants were generated to determine whether these granules are associated with the helical conformation determined from the crystallographic data. The amino acids comprising the region important for the inter-hexameric loop association ([51]YIDELASKPFYKDL[63]) were deleted or replaced by alanines (N1-del and N1-ala mutants, respectively). In addition, with the aim of evaluating whether the protonation of Lys57 is involved in helix assembly, it was mutated to arginine, another basic amino acid, or to alanine (N1-K57R and N1-K57A mutants, respectively). All of these mutants (Table 3) were fused to an additional protein–protein interaction site provided by GFP and expressed in epimastigotes. Wild-type TcNDPK1 was used as a control (N1). As shown in Fig. 7(*a*), all parasite populations expressed the same levels of fusion protein with the expected molecular weight (~40 kDa; Fig. 7*b*). Additionally, in a native PAGE Western blot the constructs had less motility than a nonhexameric mutant (P95S) previously demonstrated to affect TcNDPK1 granules

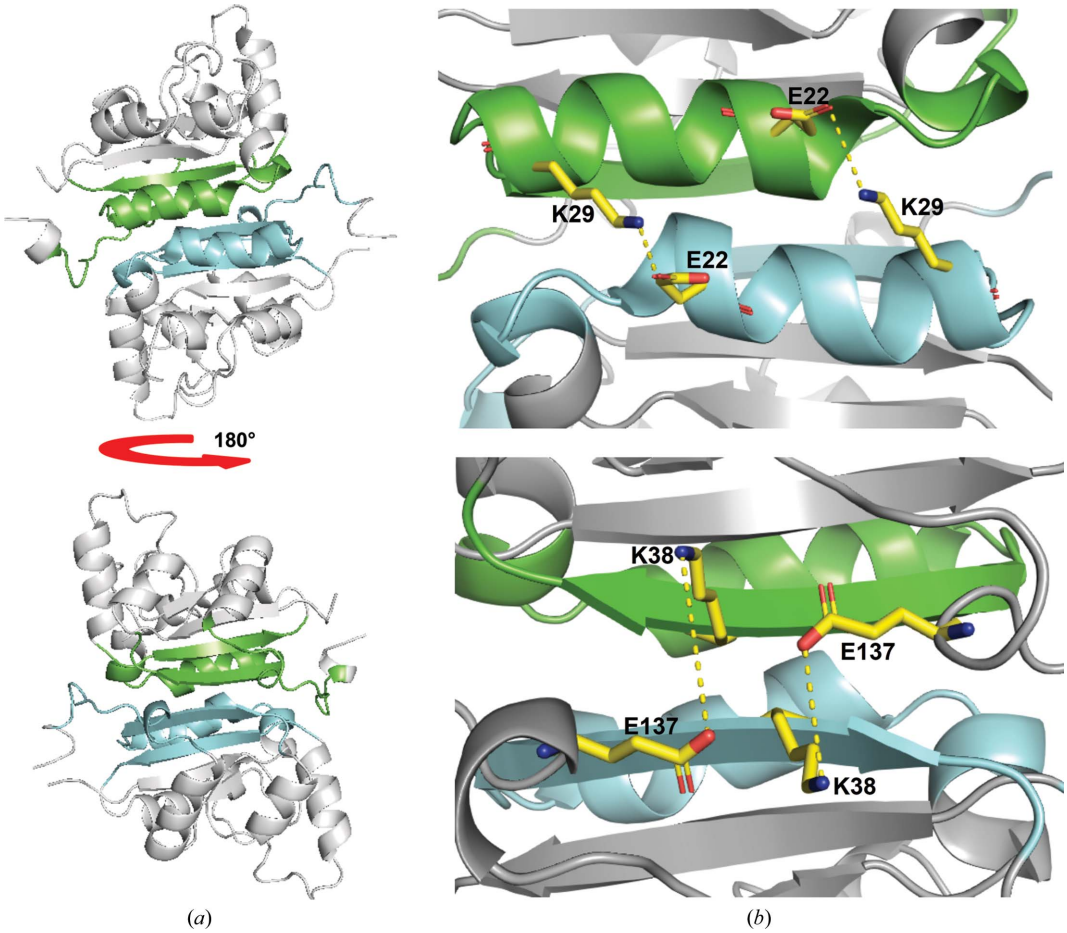


Figure 3
Dimer assembly. (*a*) The interface regions involved in the isologous dimerization are colored green and light blue. (*b*) The residues involved in ionic interactions are shown as sticks and labeled. The salt bridges Lys38–Glu317 and Glu22–Lys29 are shown.

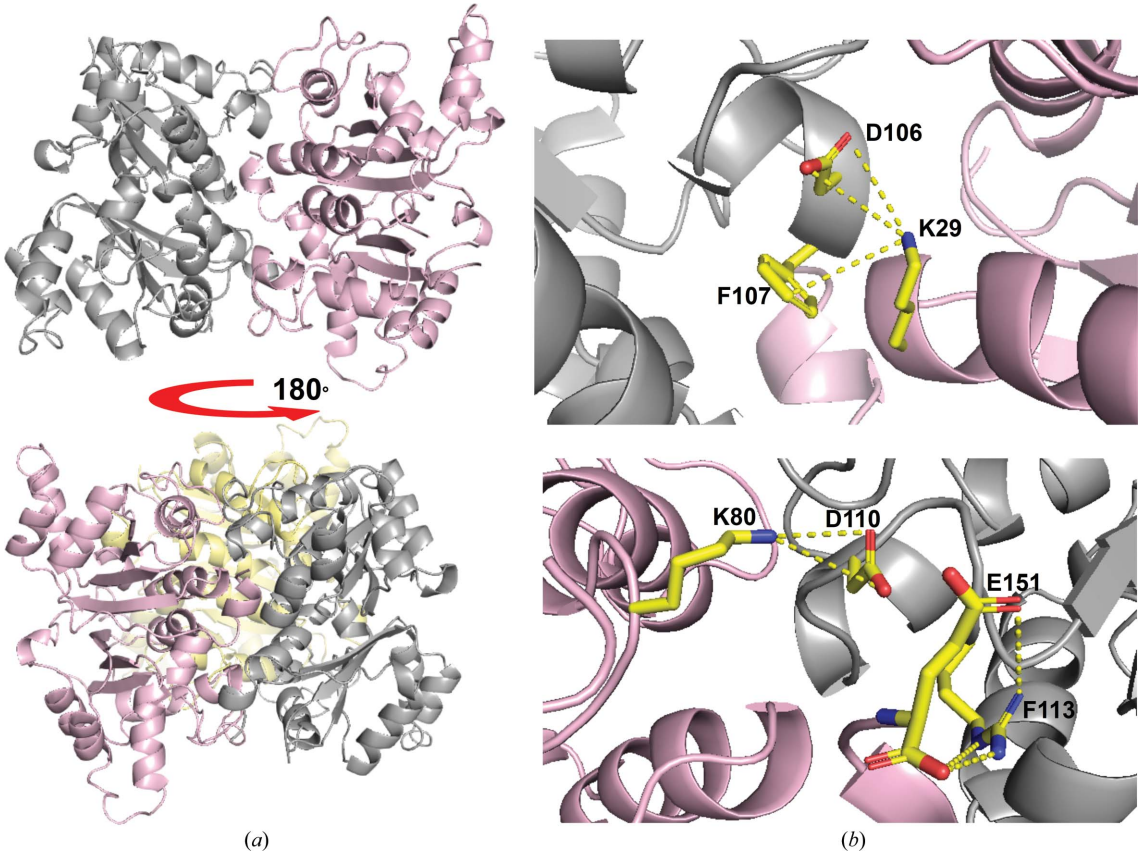


Figure 4
Hexamer assembly. (*a*) Heterologous dimer association; monomers are colored grey and magenta. (*b*) The main ionic interactions are shown as sticks and labeled. The salt bridges Aso106–Lys29, Asp110–Lys80 and Arg113–Glu151 and the cation–π interaction Lys29–Phe107 are shown.

(Pereira *et al.*, 2014), indicating that such mutations do not affect hexamerization (Fig. 7*b*). The populations were analyzed by fluorescence microscopy (Fig. 8). Parasites with soluble TcNDPK1, granules or both were observed. Granule stability was evaluated by counting the parasites containing granules and/or soluble protein in the cytosol. If the *in vitro* helix assembly is associated with the *in vivo* formation of granules, mutations would affect them and this would be reflected in the dissociation of the granules leaving enzyme in the cytosol. With regard to the presence of granules, the statistical analysis revealed that there were significant differences between the control and the N1-del and N1-ala populations: almost 100% of N1 parasites contained TcNDPK1 granules, while N1-del and N1-ala parasites reached approximately 40% (Fig. 9). Additionally, the presence of soluble TcNDPK1 in the cytosol showed significant differences between the control, N1-del, N1-ala and N1-K57A populations: while 100% of N1-del and N1-ala parasites presented soluble protein, only 16% of N1 and N1-K57R parasites and 30% of N1-K57A parasites did. Interestingly, the differences between N1-K57A and N1-K57R were statistically significant (Fig. 9). Such a difference could be attributed to a lower stability of K57A granules, since alanine is not able to fulfill the behaviour of lysine like arginine. These results indicate that this region is important for granule formation and that Lys57 is one of the main amino acids involved in the occurrence of such behavior.

Furthermore, TcNDPK1 formed filaments *in vitro* that showed a significant behavior in the presence of different nucleotides. In preliminary assays we observed that filament stability varied according to the nucleotide substrate added. In the presence of triphosphate nucleosides such as ATP, dATP or GTP the filaments disassembled quickly, while in the presence of the diphosphate nucleosides ADP or GDP few or no changes were detected over time (Supplementary Video S1).

Altogether, the evidence suggested that TcNDPK1 filaments and granules are formed by bundling of the crystallographic helical oligomers along the *z* axis.

4. Discussion

X-ray diffraction studies of *T. cruzi* nucleoside diphosphate kinase 1 (TcNDPK1), a key enzyme in the energy metabolism of the parasite, revealed a novel and quite unexpected supramolecular helical oligomer structure. The quinary structure of TcNDPK1 reveals a new level of protein complexity for NDPKs.

TcNDPK1 crystallized in space group *P*3. The symmetry of this space group is compatible with the geometry of the *in vitro* helical oligomer and the formation of the *in vivo* granules observed in *T. cruzi*. Crystallization was difficult and protracted. The crystals had a rod-like shape and were inherently unstable. They grew for a while and then started

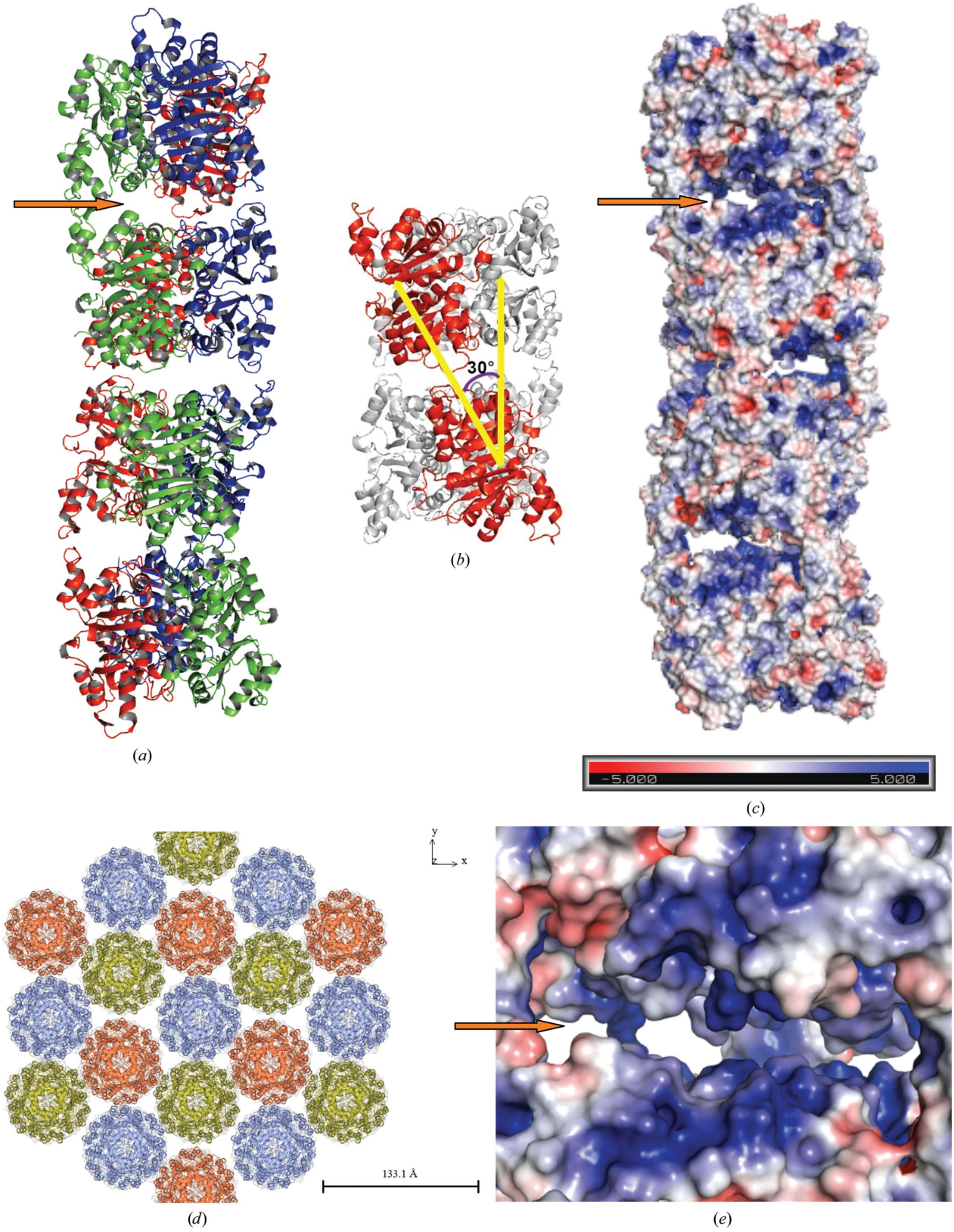


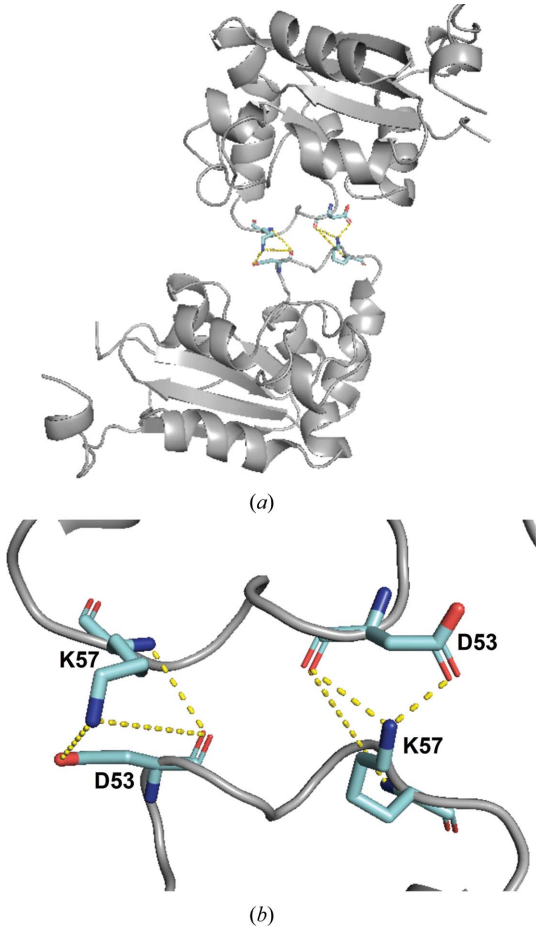
Figure 5
Quinary structure: formation of the oligomer helix. (*a*) Cartoon representation of the 24 hexamers in the asymmetric unit assembled into the left-handed oligomer helix. (*b*) Cartoon representation of the 30° hexamer rotation. (*c*) Connolly surface and electrostatic potential calculated using *APBS* and *PyMOL*. An arrow points towards the active-site cleft. (*d*) Supramolecular association of the left-handed oligomer helices generated by the trigonal crystallographic symmetry along the *z* axis to form the *in vivo* granules. (*e*) Connolly surface and electrostatic distribution of a close-up of the interhexameric interface. The arrow points to the active site. Note the highly positively charged active site for the recognition and binding of highly negative substrates.

fraying at both ends. Our efforts to improve crystal quality using standard methods for stabilizing crystals were unsuccessful. Data were collected to 3.5 Å resolution with R_{merge} = 0.34, $R_{p.i.m}$ = 0.32 and R_{meas} = 0.47. In retrospect, the failure to improve crystal quality was due to the transient dynamic characteristics of the quinary structure of TcNDPK1.

The calculated Matthews coefficient suggested 24 molecules, or four hexamers, per asymmetric unit. Unexpectedly, the calculated TcNDPK1 self-rotation function had a different point-group symmetry to *D*3. There were three extra coplanar twofold axes that could not be explained at this stage of structure solution.

The low quality of the data set combined with the inherent phase bias of the method made the molecular-replacement results very difficult to interpret. The criterion that the point-group symmetry of the solution must be compatible with the NCS observed in the self-rotation function was indispensable in choosing the real solution.

It was now clear that the extra peaks in the self-rotation function for TcNDPK1 were due to the left-handed 30° rotation of the hexamers in the asymmetric unit to form a linear helical oligomer along the *z* axis. The helical oligomers are packed parallel to the *z* axis in a bundle-like form.

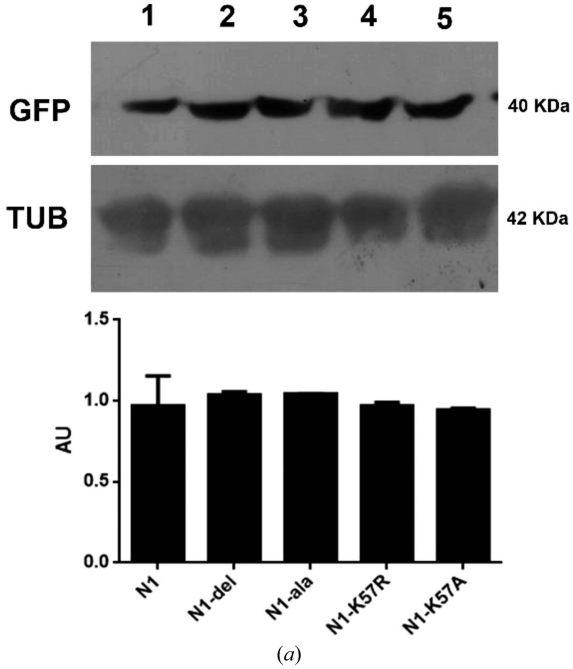


(*a*)

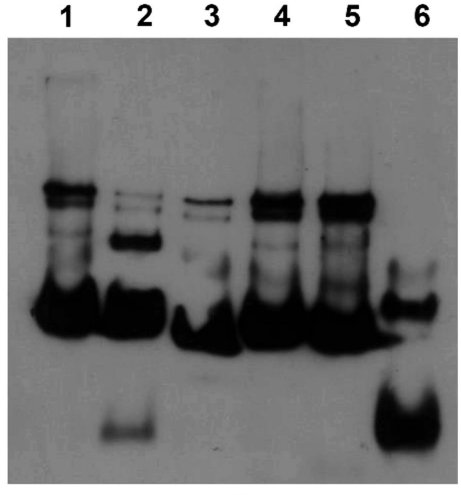
(*b*)

Figure 6
The inter-hexameric interface. (*a*) Cartoon of one of the three loops involved in inter-hexameric binding. The salt bridge involved in the binding, Asp53–Lys57, is represented as sticks. (*b*) Close-up of (*a*).

The structure solution begged the question: is the TcNDPK1 *in vitro* oligomer structure just a crystallographic artifact or does it perhaps describe the *in vivo* molecular structure of the granules previously reported in *T. cruzi* parasites by Pereira and coworkers? To answer this question different mutants were made to determine whether the *in vivo* granules were associated with the helical oligomer determined from the crystallographic data. The amino acids comprising the labile inter-hexamer loop were deleted or replaced by



(*a*)



(*b*)

Figure 7
Anti-GFP Western blot of parasites expressing TcNDPK1-GFP fusions. (*a*) SDS–PAGE and densitometry showing correct expression of the transgenes and equal levels of expression among populations. (*b*) Native PAGE of different populations showing that the mutations do not affect hexamerization. Anti-tubulin antibodies were used as a loading control. Lane 1, N1 (TcNDPK1 without mutation); lane 2; N1-del (TcNDPK1 with Tyr51–Leu63 region deleted); lane 3, N1-ala (TcNDPK1 with Tyr51–Leu63 region replaced by alanines); lane 4, N1-K57R (TcNDPK1 with Lys57 replaced by Arg); lane 5, N1-K57A (TcNDPK1 with Lys57 replaced by Ala); lane 6, N1-P95S (TcNDPK1 with Pro95 replaced by Ser in addition to deletion of the four carboxy-terminal amino acids).

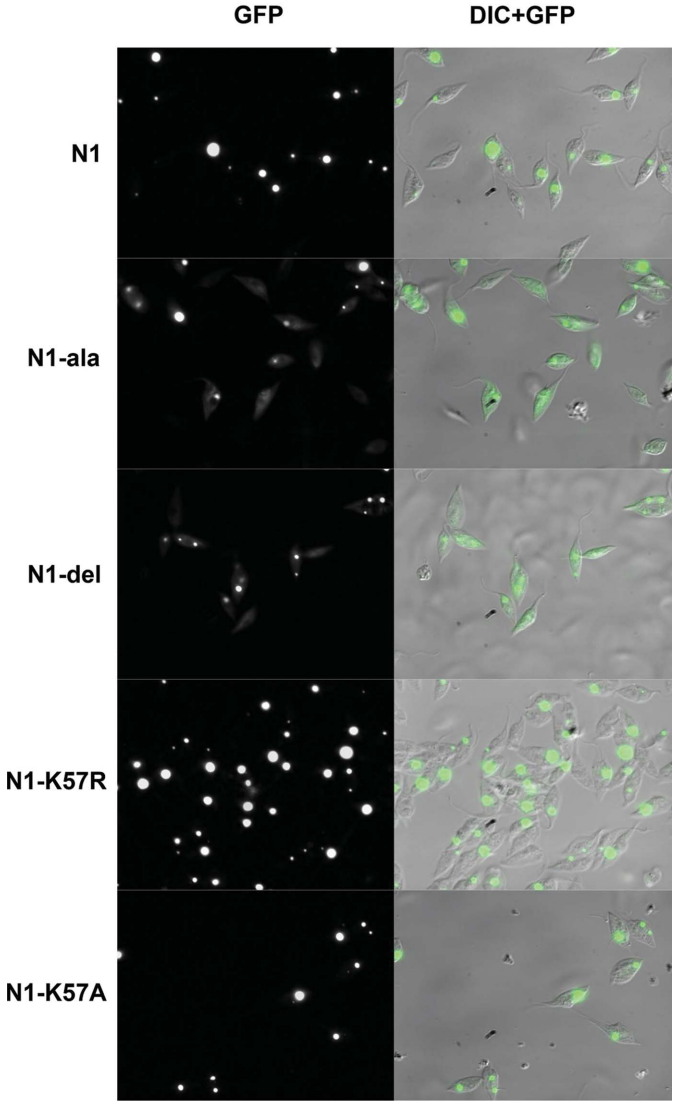


Figure 8
Fluorescence microscopy of parasites expressing the different mutations. Parasites were fixed, mounted and observed by GFP direct fluorescence. The figure shows one representative field of each parasite population. GFP: fluorescence of the different mutated TcNDPK1 proteins fused to the green fluorescence protein. DIC+GFP: merge of bright-field (DIC) and fluorescence-field (GFP) images.

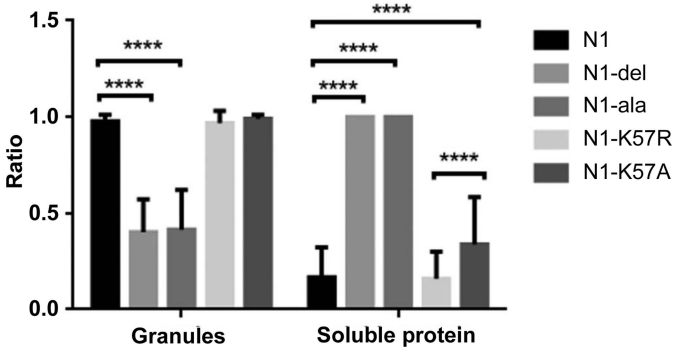


Figure 9
Statistical analysis of parasite populations. Approximately 1000 parasites were analysed for the presence of granules or soluble protein. Two-way ANOVA and *post hoc* multiple pairwise comparisons using Tukey's test were performed ($p < 0.05$). **** , $p < 0.0001$.

alanines. Also, Lys57, which is responsible for the salt bridge with Asp53 in the loop, was mutated to Ala and Arg to check its influence on helix assembly.

The results revealed that modification of the loop and disruption of the salt bridge obstruct granule formation, proving that the *in vitro* crystallographic helix is related to the granules observed *in vivo*. In other words, the granules are generated by bundling of the crystallographic helical oligomers along the *z* axis.

The three-dimensional structure of the multi-hexamer helix shows that the active site of each monomer is located in the interface between the hexamers. Additionally, the conserved residues implicated in nucleotide binding, Tyr51, Leu54, Phe59 and Tyr66, are located in the helix–loop–helix involved in inter-hexamer association. Also, Tyr51 and Leu54 participate in the hydrophobic core that stabilizes the conformation of this helix–loop–helix motif. These observations suggest that substrate binding or release would be related to the assembly and disassembly of the multihexameric complex and, as a consequence, granule assembly.

The quinary structure of TcNDPK1, a transient dynamic supramolecular helical oligomer, reveals a dynamic regulation mechanism related to the multiple functions associated with this protein. Additionally, storage and protection from degradation are consequences of active TcNDPK1 filamentation. Furthermore, the fact that the presence of triphosphate nucleosides quickly disassembles *in vitro* filaments, while the presence of diphosphate nucleosides leads to few or no changes over time, suggests that regulation of the nucleoside diphosphate–triphosphate pool is related to filament and granule formation.

Acknowledgements

We thank Dr Mariana Bentosela for statistical support. CAP and MRM are members of the career of scientific investigator, CONICET; and JAGB and CFA are Professors and Scientific Investigators of the Universidad Nacional de San Luis.

Funding information

The following funding is acknowledged: Consejo Nacional de Investigaciones Científicas y Técnicas; Agencia Nacional de Promoción Científica y Tecnológica (PICT No. 2015-0539; PICT No. 2018-1801; PICT No. 2018-1871); Universidad Nacional de San Luis (PROICO No. 9700; PROICO No. 2-0114; PROICO No. 2-0118). The work has, in part, received funding from UK Research and Innovation via the Global Challenges Research Fund under grant agreement 'A Global Network for Neglected Tropical Diseases', grant number MR/P027989/1.

References

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M.,

Cherevach, I., Chillingworth, T., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., MacLeod, A., Mooney, P. J., Moule, S., Martin, D. M. A., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M. R., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E. & El-Sayed, N. M. (2005). *Science*, **309**, 416–422.

Boissan, M., Dabernat, S., Peuchant, E., Schlattner, U., Lascu, I. & Lacombe, M.-L. (2009). *Mol. Cell. Biochem.* **329**, 51–62.

Bosnar, M. H., Bago, R. & Ćetković, H. (2009). *Mol. Cell. Biochem.* **329**, 63–71.

Bouvier, L. A., Cámara, M. M., Canepa, G. E., Miranda, M. R. & Pereira, C. A. (2013). *PLoS One*, **8**, e80217.

Camara, L. M., Bouvier, L., Reigada, C., Digirolamo, F. A., Saye, M. & Pereira, C. A. (2017). *Fol. Parasitol.* **64**, 006.

Camargo, E. P. (1964). *Rev. Inst. Med. Trop. Sao Paulo*, **6**, 93–100.

Campeotto, I., Lebedev, A., Schreurs, A., Kroon-Batenburg, L., Lowe, E., Phillips, S., Murshudov, G. N. & Pearson, A. R. (2018). *Sci. Rep.* **8**, 14876.

Castellanos-Gonzalez, A., Martinez-Traverso, G., Fishbeck, K., Nava, S. & White, A. C. Jr (2019). *Sci. Rep.* **9**, 12153.

Chen, K., Long, D. S., Lute, S. C., Levy, M. J., Brorson, K. A. & Keire, D. A. (2016). *J. Pharm. Biomed. Anal.* **128**, 398–407.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org/.

Dexheimer, T. S., Carey, S. S., Zuohe, S., Gokhale, V. M., Hu, X., Murata, L. B., Maes, E. M., Weichsel, A., Sun, D., Meuillet, E. J., Montfort, W. R. & Hurley, L. H. (2009). *Mol. Cancer Ther.* **8**, 1363–1377.

Dey, S., Pal, A., Chakrabarti, P. & Janin, J. (2010). *J. Mol. Biol.* **398**, 146–160.

Dorion, S. & Rivoal, J. (2018). *Plant. Signal. Behav.* **13**, e1475804.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Georgescauld, F., Song, Y. & Dautant, A. (2020). *Int. J. Mol. Sci.* **21**, 6779.

Gómez Barroso, J. A., Pereira, H., Miranda, M., Pereira, C., Garratt, R. C. & Aguilar, C. F. (2010). *Acta Cryst.* F**66**, 862–865.

Isupov, M. N. & Lebedev, A. A. (2008). *Acta Cryst.* D**64**, 90–98.

Kolli, B. K., Kostal, J., Zaborina, O., Chakrabarty, A. M. & Chang, K.-P. (2008). *Mol. Biochem. Parasitol.* **158**, 163–175.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Levy, E. D. & Teichmann, S. (2013). *Prog. Mol. Biol. Transl. Sci.* **117**, 25–51.

Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, J. S. (2003). *Proteins*, **50**, 437–450.

Lynch, E. M., Kollman, J. M. & Webb, B. A. (2020). *Curr. Opin. Cell Biol.* **66**, 28–33.

Mahdavi, S., Salehzadeh-Yazdi, A., Mohades, A. & Masoudi-Nejad, A. (2013). *Comput. Biol. Chem.* **47**, 16–23.

McConkey, E. H. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 3236–3240.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Miranda, M. R., Camara, M., Bouvier, L. A. & Pereira, C. A. (2011). *Mol. Biochem. Parasitol.* **177**, 152–155.

Miranda, M. R., Canepa, G. E., Bouvier, L. A. & Pereira, C. A. (2008). *Parasitology*, **135**, 1661–1666.

Monod, J., Wyman, J. & Changeux, J.-P. (1965). *J. Mol. Biol.* **12**, 88–118.

Moréra, S., LeBras, G., Lascu, I., Lacornbe, M. L., Véron, M. & Janin, J. (1994). *J. Mol. Biol.* **243**, 873–890.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). *J. Mol. Biol.* **247**, 536–540.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nooren, I. M. A. & Thornton, J. M. (2003). *J. Mol. Biol.* **325**, 991–1018.

Orengo, C., Michie, A., Jones, S., Jones, D., Swindells, M. & Thornton, J. (1997). *Structure*, **5**, 1093–1108.

Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.

Park, C. K. & Horton, N. C. (2019). *Biophys. Rev.* **11**, 927–994.

Parks, R. E. Jr & Agarwal, R. (1973). *The Enzymes*, 3rd ed., edited by P. D. Boyer, Vol. 8, pp. 307–333. New York: Academic Press.

Pereira, C. A., Reigada, C., Sayé, M., Digirolamo, F. A. & Miranda, M. R. (2014). *Exp. Parasitol.* **142**, 43–50.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.

Puts, G. S., Leonard, M. K., Pamidimukkala, N. V., Snyder, D. E. & Kaetzel, D. M. (2018). *Lab. Invest.* **98**, 211–218.

Reigada, C., Sayé, M., Girolamo, F. D., Valera-Vera, E. A., Pereira, C. A. & Miranda, M. R. (2020). *Mem. Inst. Oswaldo Cruz*, **115**, e200019.

Roversi, P., Blanc, E., Johnson, S. & Lea, S. M. (2012). *Acta Cryst.* D**68**, 418–424.

Scapin, G. (2013). *Acta Cryst.* D**69**, 2266–2275.

Simonet, J. C., Burrell, A. L., Kollman, J. M. & Peterson, J. R. (2020). *Mol. Biol. Cell*, **31**, 1201–1205.

Souza, T. A., Trindade, D. M., Tonoli, C. C., Santos, C. R., Ward, R. J., Arni, R. K., Oliveira, A. H. & Murakami, M. T. (2011). *Mol. Biosyst.* **7**, 2189–2195.

Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.

Sun, J., Singh, V., Lau, A., Stokes, R. W., Obregón-Henao, A., Orme, I. M., Wong, D., Av-Gay, Y. & Hmama, Z. (2013). *PLoS Pathog.* **9**, e1003499.

Thompson, M. C. (2017). *Methods Mol. Biol.* **1607**, 185–217.

Tina, K. G., Bhadra, R. & Srinivasan, N. (2007). *Nucleic Acids Res.* **35**, W473–W476.

Ulloa, R. M., Muschietti, J. P., Veron, M., Torres, H. N. & Tellez-Iñón, M. T. (1995). *Mol. Biochem. Parasitol.* **70**, 119–129.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.

Vieira, P. S., de Jesus Santos, A. P. & de Oliveira, A. H. C. (2016). *Protein Pept. Lett.* **23**, 99–106.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.

Wirth, A. J. & Gruebele, M. (2013). *Bioessays*, **35**, 984–993.

Yeates, T. O. (1988). *Acta Cryst.* A**44**, 142–144.

Yeates, T. O. & Yu, F. (2008). *Acta Cryst.* D**64**, 1158–1164.

Yu, H., Xiong, J., Zhang, R., Hu, X., Qiu, J., Zhang, D., Xu, X., Xin, R., He, X., Xie, W., Sheng, H., Chen, Q., Zhang, L., Rao, X. & Zhang, K. (2016). *Sci. Rep.* **6**, 28684.