Chapter 2

Investigation of the Structural Metal Dependency of *P. falciparum* Arginase with Molecular Dynamics

2.1 Introduction

2.1.1 Structure and reaction mechanism

Arginase (L-arginine-urea hydrolase, EC 3.5.3.1) catalyses the hydrolysis of (L-)arginine to (L-)ornithine and urea (Fig. 2.1). The arginase fold also includes agmatinase (Ouzounis and Kyrpides, 1994; Ahn *et al.*, 2004) which converts agmatine to putrescine and urea. Agmatine is formed by decarboxylation of arginine (arginine decarboxylase). Arginase is thus part of one of two alternative biosynthetic routes to putrescine (Fig. 1.2). Arginase and agmatinase form part of the ureohydrolase superfamily, that also includes histone de-acetylase and acetyl-polyamine amidohydrolase (Dowling *et al.*, 2008). This superfamily is considered to be a very ancient fold that was probably present in the last universal common ancestor (Ouzounis *et al.*, 2006). Arginase is thought to have evolved later from the agmatinase fold since the presence of ornithine early in evolution may have been detrimental due to mis-incorporation into proteins (Sekowska *et al.*, 2000).

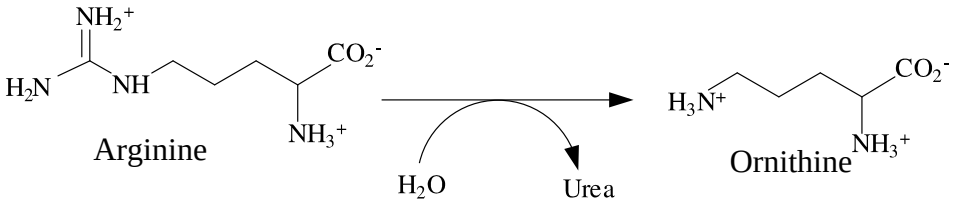


Figure 2.1: Arginase reaction

Arginase is a multimeric metallo-enzyme with a binuclear spin-coupled Mn^{2+} cluster in the active site. A number of structures have been determined for arginase. Among these are structures from the bacteria *Bacillus caldovelox* (Bewley *et al.*, 1996, 1999), human arginase I (Di Costanzo *et al.*, 2005) and II (Cama *et al.*, 2003a), and rat arginase I (Kanyo *et al.*, 1996; Scolnick *et al.*, 1997). A structure has also been deposited in the PDB for arginase from *Thermus thermophilus* (PDB ids: 2EF4, 2EF5, 2EIV). Each active site is restricted to a single α/β monomer. The monomer comprises an eight stranded parallel β-sheet packed between α-helices on either face. The β-strands are arranged in the order $\beta2 - \beta1 - \beta3 - \beta8 - \beta7 - \beta4 - \beta5 - \beta6$ (Fig. 2.2).

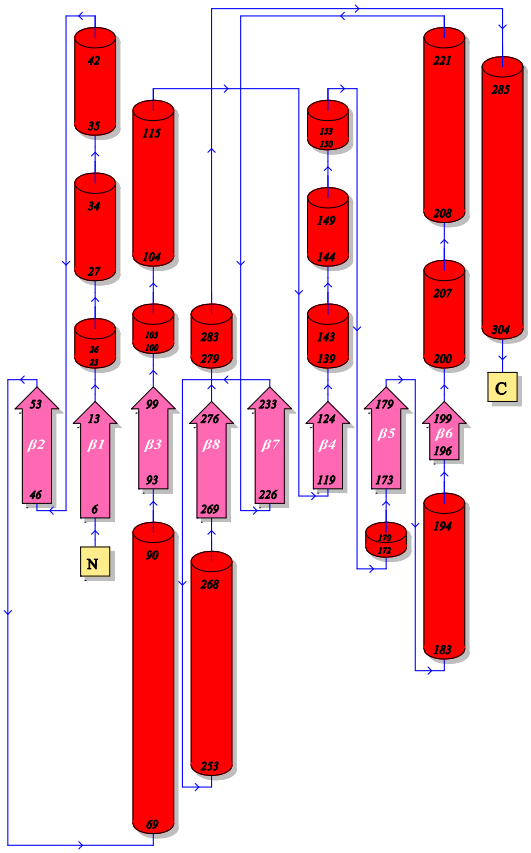


Figure 2.2: Topology of the monomer from rat arginase I (PDB id: 1RLA). Residue numbers for secondary structure element termini is included. β-sheets are numbered as indicated. Helix numbers are omitted due to the variable nature of these between species. The figure was adapted from output generated on the PDBsum server (http://www.ebi.ac.uk/pdbsum/).

The metal cluster sits in the active site in a 15 Å deep cleft at the *C*-terminal end of the β-sheet. The Mn^{2+} atoms are about 3.3 Å apart and are bridged by a solvent molecule (Kanyo *et al.*, 1996; Bewley *et al.*, 1999). Mn_A^{2+} is the more inaccessible of the two ions and was initially reported to be co-ordinated by five ligands in rat arginase I with a distorted square pyramidal geometry. The co-ordinating ligands comprise one His residue, three monodentate Asp residues and the bridging solvent. The second ion (Mn_B^{2+}) was reported to be co-ordinated by six ligands in a distorted octahedral geometry (Scolnick *et al.*, 1997), including one His, two monodentate Asp residues, one bidentate Asp residue and the bridging solvent. The two Mn^{2+} ions are bridged by three ligands: one bidentate Asp, one monodentate Asp and the bridging solvent. X-ray absorption spectroscopy suggested both ions are co-ordinated octahedrally (Stemmler *et al.*, 1997), however, with an extra solvent molecule binding to Mn_A^{2+}. This was confirmed by the bacterial structure (Bewley *et al.*, 1999) and re-examination of the original rat I arginase structure (Kanyo *et al.*, 1996) (Fig.

2.3). The correspondence of the rat I, human II and bacterial Mn^{2+}co-ordinating residues is given in Table 2.1. The requirement of these ligands for enzyme activity have been confirmed by a number of mutagenesis studies. Mutation of His 101 and His 126 to Asn in rat arginase I compromise activity (Cavalli *et al.*, 1994). The first of these disrupts Mn_A^{2+} and thermostability (Scolnick *et al.*, 1997). Furthermore, the mutations Asp 233 Ala/Cys, Asp 128 Glu/Asn/Ala, Asp 234 Glu/Ala/His and His 101 Glu all compromise activity and Mn^{2+} binding (Cama *et al.*, 2003b).

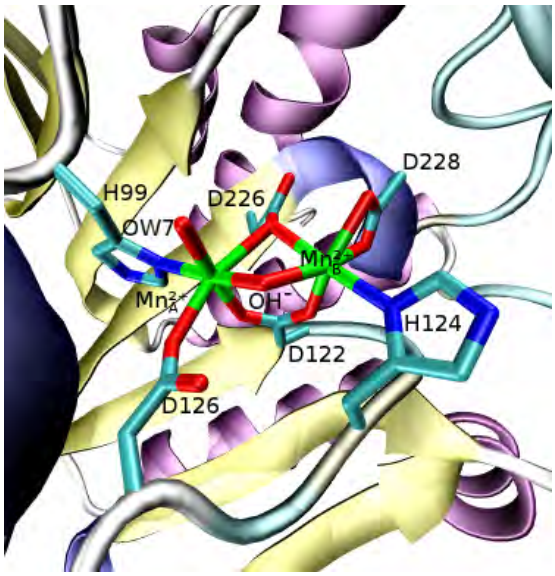


Figure 2.3: Co-ordination pattern of Mn^{2+} in bacterial arginase (PDB id: 2CEV). Both Mn^{2+} ions display octahedral co-ordination, however, in some cases Mn_A^{2+} shows square-pyrimadal co-ordination (The water molecule OW7 is absent).

Table 2.1: Correspondence of Mn^{2+} co-ordinating and other active site residues in rat I-, human II- and bacterial- (*Bacillus caldevelox*) arginase.

Rat I	Human II	Bacterial	Rat I	Human II	Bacterial
His 100	His 120	His 99	His 141	His 160	His 139
Asp 124	Asp 143	Asp 122	Asp 232	Asp 225	Asp 226
His 126	His 145	His 124	Asp 234	Asp 227	Asp 228
Asp 128	Asp 147	Asp 126			

In the mechanism proposed by Di Costanzo *et al.* (2005) the metal bridging solvent is an activated hydroxyl which attacks the guanidinium carbon of arginine, followed by collapse of the tetrahedral intermediate to release ornithine and urea (Fig. 2.4). However, *ab initio* modeling of the active site suggests that the bridging solvent may be a neutral water molecule instead (Ivanov and Klein, 2004). The high resolution structure of human arginase I reveals that His 141 is protonated and most probably acts as a general acid to

protonate the leaving group of L-ornithine. The catalytic role of this residue is confirmed by mutagenesis to Asn (Cavalli *et al.*, 1994) and Phe (Carvajal *et al.*, 1999a) which substantially reduce activity but have no effect on Mn^{2+}binding or K_m. The mechanism is confirmed by the binding of various inhibitors that are expected to mimic the tetrahedral transition state. Among these are the borate ion $B(OH)_4^-$(Carvajal *et al.*, 1999b), inhibitors with the trigonal borono ($-B(OH)_2$) moiety which is attacked by the nucleophillic OH^- to form a tetrahedral intermediate ($-B(OH)_3^-$) and sulphonamides (Cox *et al.*, 2001; Cama *et al.*, 2003a,c, 2004). The metal-bridging by the carbonyl moiety of urea is confirmed by a similar binding mode for the C = S group of thiosemicarbazide (Di Costanzo *et al.*, 2005).

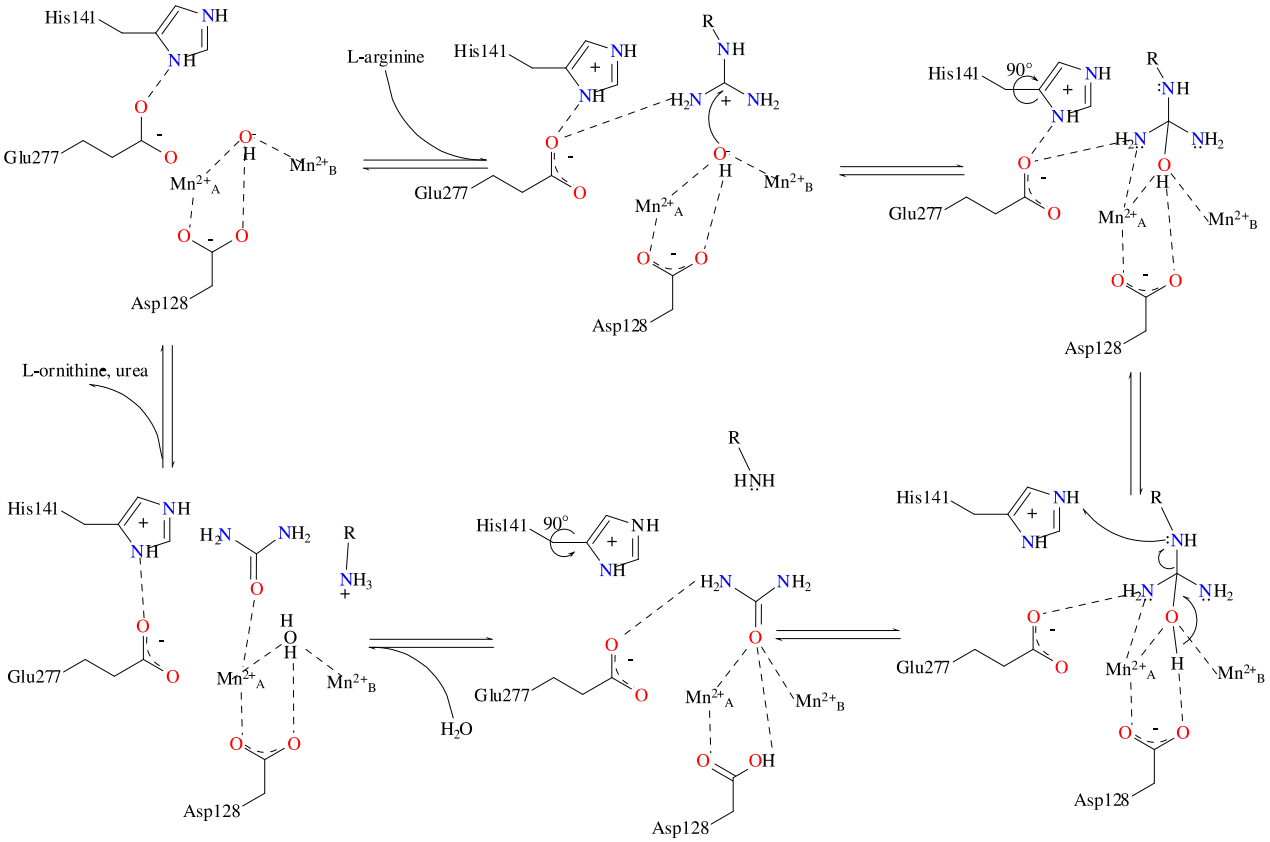


Figure 2.4: General reaction mechanism of arginase, adapted from Di Costanzo *et al.* (2005). Residue numbers correspond to rat I arginase. The activated hydroxyl attacks the guanidinium carbon of arginine, followed by collapse of the tetrahedral intermediate to release ornithine and urea. See main text for further details.

2.1.2 Quaternary structure

In most recent literature arginase is observed within trimeric or hexameric complexes. Carvajal *et al.* (1971) did report human arginase I as being tetrameric, however, this is probably because the molecular mass of arginase was underestimated to be 30 kDa, as opposed to the true molecular mass of 34.7 kDa (Di Costanzo *et al.*, 2007a). Eukaryotic arginases are observed to be trimeric, whereas bacterial arginases are hexameric (Jenkinson *et al.*, 1996). Furthermore, in *S. cerevisiae* trimeric arginase forms a hetero-hexameric complex

with trimeric ornithine transcarbamoylase. Ornithine transcarbamoylase catalyses the first committed step in arginine biosynthesis, and occurs in the cytosol together with arginase in yeast. The presence of both these enzymes could potentially keep urea in a futile cycle leading to hydrolysis of ATP, however, when arginase complexes with and inactivates ornithine transcarbamoylase (Bechet and Wiame, 1965; Messenguy and Wiame, 1969).

In human arginase I the main structural determinant for oligomerisation was found to be an inter-monomer salt-bridge between Glu 256 and Arg 255. Mutating Glu 246 to Gln abolished trimer formation, with no effect on kinetic properties (Sabio *et al.*, 2001). This salt-bridge has been observed in all arginase structures so far, and is likely to be necessary in all arginases for oligomerisation.

In the first structure of rat arginase I Kanyo *et al.* (1996) reported that mostly inter-monomer contact was made by a novel S-shaped motif at the *C*-terminus. This was tested in human arginase I by deletion of fourteen residues following Arg 308 (= rat I Arg 308) which apparently nucleates the interactions of the S-shaped motif by forming a salt bridge with Asp 204. The *C*-terminus was not found to be required for oligomerisation of human arginase I. Furthermore, treatment with high salt concentrations did not disrupt the oligomer, suggesting that Arg 308 is also not required for oligomerisation. The kinetic properties were also unaltered (Mora *et al.*, 2000). The role of Arg 308 was later contradicted somewhat by Lavulo *et al.* (2001), where mutating Arg 308 to Ala, Glu or Lys in rat arginase I abolished trimer formation. These mutations were also found to compromise enzyme activity where k_{cat} was 33% - 41% of the wild type. The S-shaped tail is conserved in multiple isoforms of eukaryotic arginase (Lavulo *et al.*, 2001). In bacterial arginase from *Bacillus caldevelox* Arg 308 aligns with Leu instead. Furthermore, at the inter-monomer interface a guanidinium ion is observed at the corresponding position for rat I Arg 308 when crystals are grown in guanidinium hydrochloride (Fig. 2.5). This guanidinium may be replaced by free arginine, depending on the crystallisation conditions (Bewley *et al.*, 1999).

Previously, a relationship between active site metal binding and tertiary and quaternary structure in arginase has been reported. Carvajal *et al.* (1971) have reported that human arginase I requires Mn^{2+} in order to form oligomers. However, active monomers immobilised on nylon could be formed by incubation with Mn^{2+} (Carvajal *et al.*, 1977, 1978, 1982). Arginase from *Saccharomyces cervisiae* was initially reported to contain a weakly bound $Mn^{2+}$and a strongly bound Zn^{2+} ion. Removal of the strongly bound ion caused dissociation of the trimer and loss of activity. *S. cerevisiae* arginase requires a heating step at 45 °C in the presence of $Mn^{2+}$to become active. Removal of both the weakly and strongly bound ions caused a decrease in T_m from 57 °C to 54 °C and 38 °C, respectively (Green *et al.*, 1990, 1991).

Thus, in mammals substantially active monomers are still formed by mutations that abolish trimer formation (Lavulo *et al.*, 2001; Sabio *et al.*, 2001). Furthermore, mutations disrupting metal binding has no reported effect on the quaternary structure (Cavalli *et al.*, 1994; Scolnick *et al.*, 1997; Cama *et al.*, 2003b). In contrast, a stronger relationship between

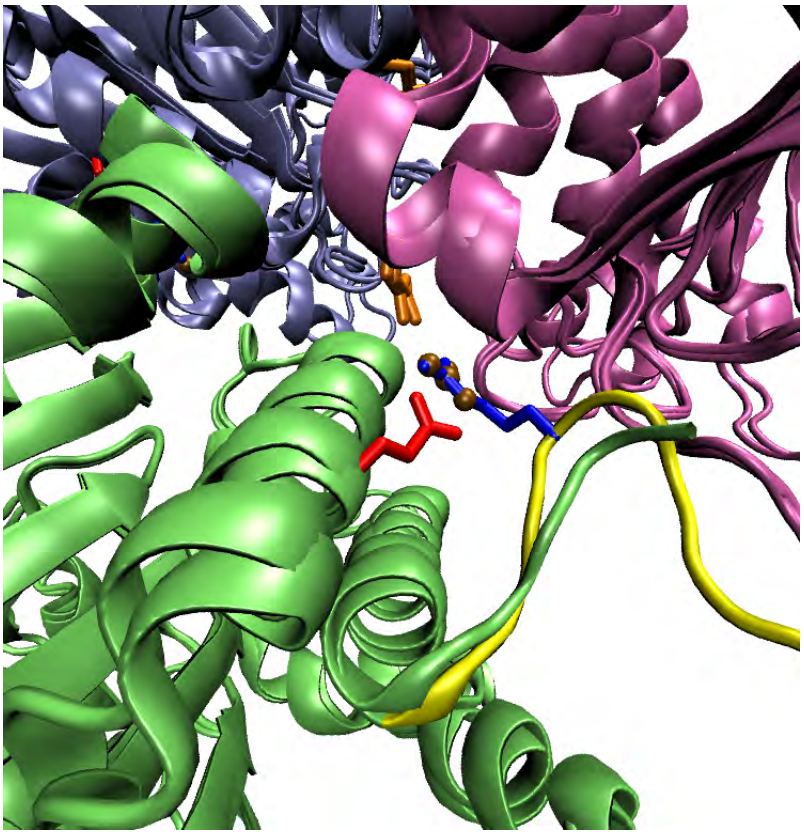


Figure 2.5: Inter-monomer interactions in rat arginase I (PDB id: 1RLA) and *B. caldovelox* arginase. Monomers are in green, light-blue and mauve. The S-shaped tail of rat arginase I is indicated in yellow. Asp204/199 (rat/*B. caldovelox*) is indicated in orange, and Glu256 of *B. caldovelox* in red. Arg308 of rat arginase I is indicated in blue, and guanidinum from *B. caldovelox* in brown. The Arg308 interaction from rat arginase I is replicated by the guanidinium bridging between Asp199 and Glu256 in *B. caldovelox*

metal binding and quarternary structure has been observed in *P. falciparum* (Müller *et al.*, 2005). Mutation of *Pf* Arg Glu 347 to Gln causes dissociation of the trimer as well as inactivation of the enzyme by 90%. This contrasts with human arginase I, where Glu 256 Gln is fully active. Furthermore, mutations to the active site that compromise activity also cause the trimer to dissociate (Table 2.2). This has yet to be reported in other species.

Table 2.2: Properties of mutated *P. falciparum* arginase in comparison to the wild type enzyme.

Name	Activity (% of wild type)	Structure
Wild type	100	Trimer
His 193 Ala	Inactive	Trimer
His 218 Ala	Inactive	Monomer
His 233 Ala	Inactive	Monomer
Asp 216 Ala	Inactive	Monomer
Asp 220 Ala	Inactive	Monomer
Asp 323/325 Ala	Inactive	Monomer
Glu 347 Gln	10	Monomer

2.1.3 Metabolic functions

In mammals, two isoforms of arginase have been identified. Arginase I is cytosolic and largely hepatic where it catalyses the final step of the urea cycle (Spector *et al.*, 1982, 1985). Non-hepatic arginase I is likely to be involved in the regulation of NO synthesis in immune tissue. Inducible NO synthase (iNOS) and arginase both compete for arginine as a substrate. These are reciprocally regulated by Th1/Th2 (Munder *et al.*, 1999), and arginase is inhibited by NOHA (N^{G}-hydroxy-L-arginine), an intermediate of iNOS (Hecker *et al.*, 1995). Arginase II is non-hepatic, and occurs in the mitochondrial matrix (Grody *et al.*, 1987; Gotoh *et al.*, 1996; Cox *et al.*, 2001). Arginase II is involved in homoeostasis of ornithine for production of proline and glutamate (Morris, 2002) and is also potentially involved in regulating NO biosynthesis (Li *et al.*, 2001; Cama *et al.*, 2003a). Bacterial arginases have a single isoform, while more than one exists in yeast (Borkovich and Weiss, 1987). In the yeast *Nuerospora crassa* two isoforms have been reported from a single gene under control of tandem promoters. Yeast arginase has been implicated in glutamate accumulation during germination and asexual spore development (Marathe *et al.*, 1998; Turner and Weiss, 2006). Arginase in *P. falciparum* is involved in the biosynthesis of the polyamine putrescine and is likely to be the sole biosynthetic route to ornithine, since no agmatinase has been identified in the parasite (Müller *et al.*, 2005).

2.2 Aims

P. falciparum arginase exhibits an interesting metal/trimer dependency compared to the human host. Mutations that abolish metal binding or removal of the metal ions by dialysis and EDTA chelation causes dissociation of the trimer into inactive monomers. Conversely, mutations that abolish inter-monomer interactions which are far from the active site result in inactive monomers (Müller *et al.*, 2005). This host-parasite difference may thus provide a novel non-active site based strategy for inhibiting *P. falciparum* arginase. Thus the mechanism of this structural dependency was investigated by homology modeling and molecular

dynamics, in order to establish an *in silico* system for exploiting this dependency.

Specifically, model systems of *P. falciparum* arginase were simulated with and without the active site metal as well as various mutants predicted to affect trimer interaction. The effects of these mutations were also tested experimentally on recombinantly expressed protein.

2.3 Methods

2.3.1 Sequence alignments

Reference multiple alignments were generated using CLUSTALW 1.82 and the FUGUE (Shi *et al.*, 2001) server. Default parameters were used in both cases. The CLUSTALW alignment included Eukaryotic Arginase types I and II and bacterial Arginases. The following sequences were used for the CLUSTALW alignment (Entrez accession number are given in brackets for non-*Plasmodium* species, PlasmoDB reference numbers are used for *Plasmodium* sequences): *Arabidopsis thaliana* (P46637, Q9ZPF5), *Schizosaccharomyces pombe* (P37818, Q10066), *Xenopus laevis* (Q91553, Q91554, Q91555, P30759), *Homo sapiens* (P78540, P05089), *Mus musculus* (O08691, Q61176), *Rattus norvegicus* (O08701, P07824), *Agrobacterium tumefaciens* (P14012), *Bacillus caldovelox* (P53608), *Bacillus subtilis* (P39138), *Brucella melitensis* (Q59174), *Coccidioides immitis* (P40906), *Emericella nidulans* (Q12611), *Neurospora crassa* (P33280), *Rana catesbeiana* (P49900), *Glycine max* (O49046), *Staphylococcus aureus* (P60086), *Saccharomyces cerevisiae* (P00812), *Plasmodium knowlesi* (PKH_070380), *Plasmodium vivax* (Pv098770), *Plasmodium falciparum* (PFI0320w), *Plasmodium yoelii* (PY03443), *Plasmodium berghei* (PB000787.03.0). Sequences for *P. knowlesi*, *P. vivax*, *P. yoelii* and *P. berghei* were obtained by using BLAST as provided on the plasmodb.org website with the *P. falciparum* sequence as query. Although the reference alignments were often highly redundant, all sequences were retained to offset the bias of including five *Plasmodium* sequences.

2.3.2 Homology modeling

MODELLER 7v7 and 8v0 were used to build homology models. Trimeric models were constructed on the rat arginase I (PDB id: 1RLA[abc], Kanyo *et al.* (1996)), human arginase II (PDB id: 1PQ3[abc], Cama *et al.* (2003a)) and *Bacillus caldovelox* (PDB id: 1CEV[abc], Bewley *et al.* (1999)) templates. Super-imposable monomers were constructed by imposing symmetry restraints on the internal co-ordinates of all atoms during the model building process (MODELLER 8 only). Models were built using very slow refinement. The effect of various sequence alignments was determined by generating multiple models with different random number seeds and monitoring the effect on the number of residues in disallowed regions of the Ramachandran plot and on the overall G-factor score from PROCHECK. Problem areas were identified as residues that frequently fell in disallowed regions. Models that minimised residues with poor phi/psi values and maximised the G-factor were used for molecular dynamics.

Three alignments were used to generate homology models. The first of these (a1, Fig. 2.6) was initiated with a CLUSTALW alignment, followed by manual adjustment with reference to existing crystal structures. The second alignment was obtained using the FUGUE structural alignment server, followed by manual adjustments (a2, Fig. 2.7). Two parasite-specific inserts occur in *Pf* Arg. The third alignment (a3) was created by moving the position of insert 2 in a1 to that of a2 (one residue downstream). In all cases the full trimer was modelled. In a2 and a3 the first 16 residues were excluded to avoid the uncertainties involved with modeling long stretches *ab initio*.

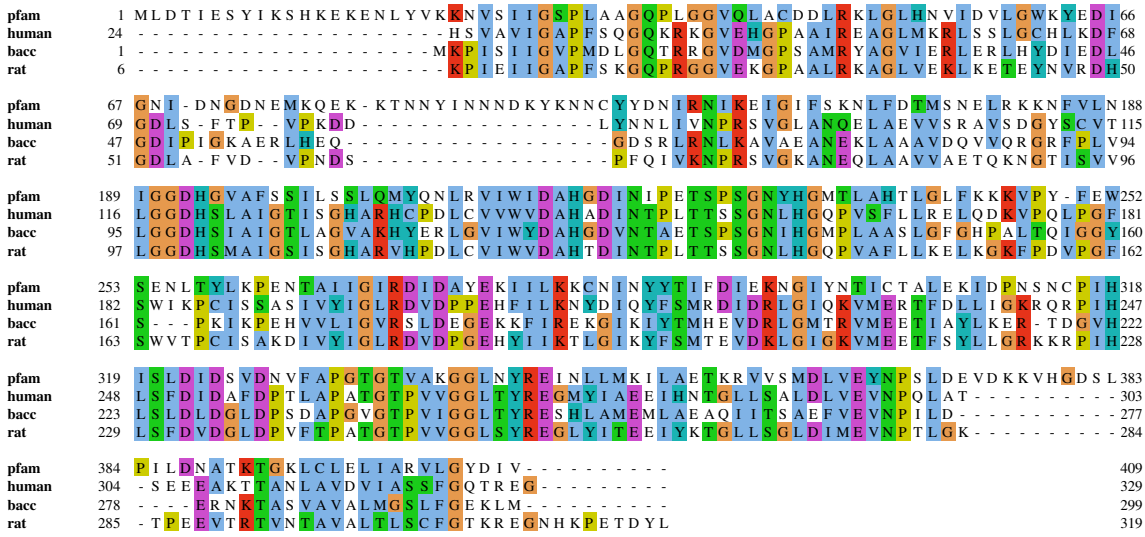


Figure 2.6: Alignment 1 (a1). Constructed with CLUSTALW multiple alignment as reference. Amino acid similarity is indicated using the CLUSTAL (Thompson *et al.*, 1997) colour scheme.

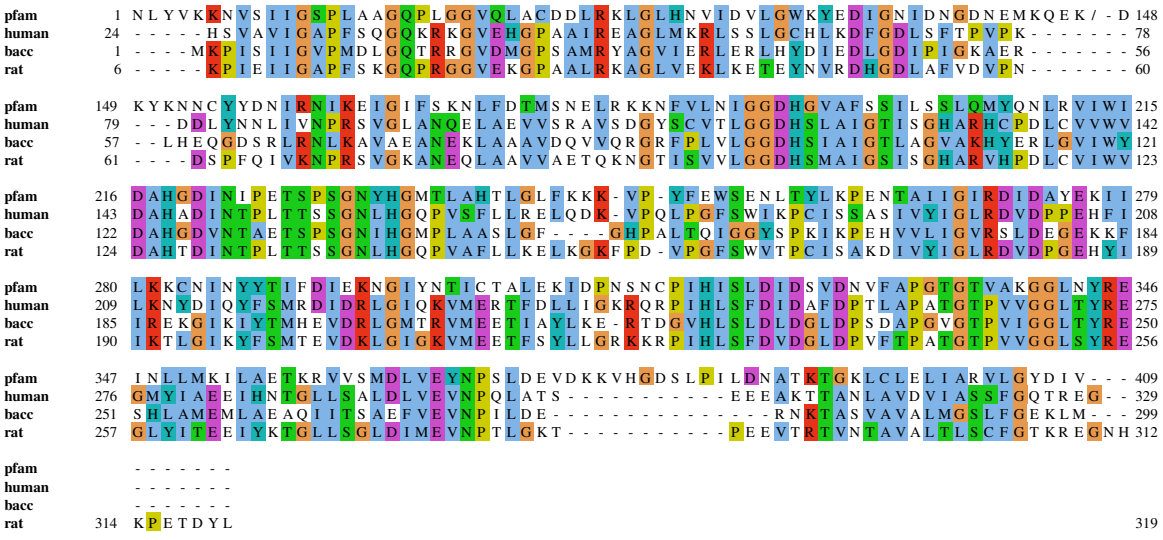


Figure 2.7: Alignment 2 (a2). Constructed with FUGUE alignment as reference. In alignment 3, insert 2 of this alignment is as in alignment 1. Amino acid similarity is indicated using the CLUSTAL (Thompson *et al.*, 1997) colour scheme.

2.3.3 Molecular dynamics

The methodology followed was essentially an exploratory approach as well as a learning process. Different protocols and software were employed in order to find the most appropriate *in silico* system for the hardware available, at the same time producing results that agree with previous experiments, especially with regards to known protein-protein interactions.

2.3.3.1 Simulations with CHARMM

System setup: Hydrogen atoms were added automatically using CHARMM (29B2 or 32B1). The histidine protonation scheme adopted is based on known requirements for catalytic activity and co-ordination of metal atoms in the active site. Thus His 193 and His 218 were protonated on N_ϵ. All other histidines were protonated on N_δ. Glu, Asp and Lys residues were charged. Since the current CHARMM forcefields are not parameterised for Mn^{2+}, the Mg^{2+} ion was used instead. It was thus assumed that any effects of the metal on trimer formation were largely electrostatic in nature (discussed below).

Long range interactions: For CHARMM simulations the same scheme was used to calculate non-bonded interactions (van der Waals and electrostatic) during minimisation and molecular dynamics. Typically, interactions were calculated between all atom pairs, except atoms connected by three or less covalent bonds (1-2, 1-3, 1-4 interactions). A constant dielectric was used, to simulate solvent attenuation of both non-bonded interactions. Non-bonded interactions were shifted to zero at from 8 Å to a cut-off of 12 Å (beyond which long-range interactions are ignored) to prevent the introduction of numerical errors.

Heating and equilibration:

Solvated shell system: An early homology model (methods not shown) was used to create a shell solvated trimer. The protein trimer was first minimised for 20 steps with steepest descent (SD). The positions of the atoms were restrained with harmonic force constraints of 20 ($\times 2 \times$MW $kcal.mol.\AA^2$). This was followed by minimisation with conjugate gradient (CJ) in various stages with a decrease in the force constant on the harmonic restraints (10: 200 steps, 5: 200 steps, 0: 100 steps). The system was then explicitly solvated with the TIP3 water model in a shell of 10 Å (55602 atoms). The whole system was then minimised with a harmonic force constant of 50 on the solute (protein) for 50 steps with steepest descent and 100 steps with conjugate gradient. The solvent was then heated (force constant 50 on solute) from 60 K to 298 K in 10 K increments every 500 steps using velocity scaling. The SHAKE algorithm was used to restrain hydrogen atom movement. The timestep was 1 fs and the run lasted 30 000 steps (30 ps).

***In vacuo* trimer - early model:** The same model as used for the solvated shell system was used for molecular dynamics *in vacuo*. The system was then minimised in various stages

as previously described. Harmonic restraints were applied to all atoms. The stages were as follows: 20 steps SD - force constant 20, 200 steps CJ - force constants 10, 200 steps CJ - force constants 5, 100 steps CJ - force constants 0. The system was then heated from 58 K to 298 K in 10 K increments every 500 steps. The SHAKE algorithm was used to restrain hydrogen atom movement. The timestep was 1 fs and the run lasted 30 000 steps (30 ps).

***In vacuo* trimer - models based on a1 and a2 alignments:** It was suspected that mistakes in the alignment used for the early model (results not shown) were resulting in poor models, and thus poor MD. Therefore, new alignments were used as described above (2.3.2) to rebuild the trimers on three different templates (the alignment used for the early model was the same as a1, however only included the human template). Also, this time the effect of imposing symmetry on internal restraints was monitored.

The system was minimised and equilibrated as described above (2.3.3.1), except that the equilibration time varied between 30 000 and 90 000 steps. The 30 000 step runs were for models based on the a1 alignment with symmetry imposed, and for the FUGUE based a2 alignment without symmetry. The 90 000 step runs were for models based on the a2 alignment with symmetry imposed, using separate runs with and without the active site Mg^{2+} atoms.

Sampling: For the production run the system was then simulated for 1 to 10 ns (10 000 000 steps) at constant energy. Sampling was only pursued for systems equilibrated for 90 000 steps.

Hardware: CHARMM was run on a cluster of 64 Pentium IV processors running Gentoo Linux, with GigaBit Ethernet interconnect at the University of Pretoria.

2.3.3.2 *In vacuo* simulations with NAMD

Although CHARMM was initially used for molecular dynamics, it was found to be too slow on the hardware that was initially available (Pentium IV Beowulf cluster with GigaBit Ethernet). NAMD 2.6 was therefore used instead because of being faster and scaling better with more processors.

Long range interactions: A cut-off of 12 Å was used for electrostatic and van der Waals interactions. These interactions were gradually tapered to 0 from 8 to 12 Å (switching function). All pairwise interactions more than 3 covalent bonds apart were included, as for the CHARMM simulations. The same non-bonded parameters were used for minimisation and dynamics.

Minimisation: For adding hydrogens CHARMM was still used. Hydrogens were added as for the CHARMM simulations: all His residues were again N_δ protonated, except for 193 and 218 which were N_ϵ protonated. The system was first minimised for 450 steps.

Dynamics: The system was heated from 60 K to 310 K in 10 K increments every 500 steps with velocity reassignment. Equilibration was for a total of 200 000 (200 ps) steps of 1 fs each. For the production run the system simulated at constant energy for 1.5 ns (with Mg^{2+}) or 12 ns (without Mg^{2+}).

Hardware: NAMD was run on a cluster of 64 Pentium IV processors running Gentoo Linux, with GigaBit Ethernet interconnect at the University of Pretoria.

2.3.3.3 Solvated simulations with NAMD - NP sampling

The following parameters were used to simulate the trimer in an isobaric (NP) ensemble.

Long range interactions: Non-bonded interactions were shifted to zero from 10 Å to a cut-off of 12 Å. All non-bonded interactions connected by more than four covalent bonds were included. The trimer complex was first minimised *in vacuo* for 450 steps.

All steps including the complete system with solvent and ions were simulated with periodic boundary conditions using Particle Mesh Ewald (PME) Sums for the electrostatic calculations. A typical cell size was about 118x Å × 113y Å × 79z Å. A PME sub-angstrom grid size of 120 × 120 × 90 points was used.

System setup: Hydrogen atoms were added automatically using CHARMM 32B1. The histidine protonation scheme adopted is based on the requirements for catalytic activity and co-ordination of metal atoms in the active site. Thus His 193 and His 218 which were protonated on N_ϵ, and His 233 which was protonated on both N-atoms. The double protonation of His 233 is in accordance with the high resolution (1.29 Å) crystal structure of human arginase I (PDB id: 2AEB, Di Costanzo *et al.*, 2005), for which hydrogen positions were also determined, as well as previous speculation concerning activity. All remaining Histidines were protonated in the N_δ atom. Glu, Asp and Lys residues were charged. Since the current CHARMM forcefields are not parametrised for Mn^{2+}, the Mg^{2+} ion was used instead. It was thus assumed that any effects of the metal on trimer formation were largely electrostatic in nature.

The system was explicitly solvated using the SOLVATE plugin of VMD using the TIP3 water model. The protein was padded with solvent for 12 Å in the x- and y-directions (xy being coplanar with the trimer), and 10 Å in the z-direction. NaCl counter ions were added with the AUTOIONIZE from VMD plugin to an ionic strength of 50 mM with charge balancing to create a net system charge of zero. The system with metals (Mg^{2+}) included 14 Na^+ ions and 11 Cl^- ions, while the system without metal finally contained 20 Na^+ ions and 5 Cl^--ions. Assuming a total protein volume of 20% of the cell size (118 Å × 113 Å × 79 Å), 25 ions gives a concentration of 50.3 mM.

$$\frac{25 \text{ ions}}{0.8 \times 117 \times 10^{-10}\, m \times 112 \times 10^{-10}\, m \times 78.7 \times 10^{-10}\, m \times 1000 \frac{l}{m^3}} \times \frac{1\, mol}{6.02214179 \times 10^{23} \text{ ions}} = 50.3\, mM$$

It was realised later after the lengthy production runs that the program uses the concentration parameter to specify total ion concentration, and therefore the amount of each ion is less. However, 50 mM is a standard ionic concentration used in the molecular dynamics literature.

Heating: The solvent of the entire system was first minimised for 2000 steps, followed by the solvent and non-backbone atoms (backbone = C,N,O,C_α) for another 2000 steps. The entire system was then heated for 20 000 steps (20 ps) from 60 K at 10 K increments every 500 steps using velocity reassignment. All atoms were then minimised for 2000 steps, followed by another heating step as described, but for 200 ps. During heating timesteps were 1 fs and a Langevin piston (piston period 100 fs, piston decay 50 fs) was used to maintain pressure at 1 atm.

Equilibration and sampling: A Langevin piston was used to maintain pressure at 1 atm (NP ensemble) with a piston period of 200 fs and a piston decay of 100 fs, whereas temperature was left to fluctuate. Total run lengths for this stage were 20 ns.

Hardware: NAMD was run on a cluster of 64 Pentium IV processors running Gentoo Linux, with GigaBit Ethernet interconnect at the University of Pretoria.

2.3.3.4 Solvated simulations with NAMD - NPT sampling

Long range interactions and system setup were the same as for the NP simulations. The following parameters were used to simulate the trimer in an isobaric-isothermal ensemble (NPT).

Heating: Two different heating protocols were used. The first protocol (NPT_1) is the same as that used for the NP ensemble. In the second protocol (NPT_2) all steps were the same as for the first, except that during the first heating only solvent atoms were allowed to move. During heating timesteps were 1 fs and a Langevin piston (piston period 100 fs, piston decay 50 fs) was used to maintain pressure at 1 atm.

It was initially intended to use the second protocol, but the first was used instead because of a typographical mistake in the input. During heating and minimisation, the general procedure is to gradually prepare the system for long dynamics runs at high temperature, since starting a simulation at the desired temperature directly, can result in numerical errors. For this, a number of different protocols are possible, thus it is not expected that using this slightly different protocol will produce vastly differing results. After gaining access to faster hardware (described below), it became feasible to repeat the molecular dynamics run using the slightly altered protocol.

Equilibration and sampling: A Langevin piston was used to maintain pressure at 1 atm with a piston period of 200 fs and a piston decay of 100 fs, and Langevin dynamics was used

to also maintain temperature at 310 K (NPT ensemble) with a damping constant of 5 ps^{-1}. Total run lengths for this stage were in the order of 20 - 50 ns.

Hardware: A number of Linux clusters were used for molecular dynamics simulations. These include a 64 processor Gentoo Linux cluster of Pentium IV processors (University of Pretoria), Clusters of Xeon, Opteron or Itanium2 processors (BMIC, Meraka Institute, CSIR) running Scientific Linux. In both cases the interconnect comprises GigaBit Ethernet. Temporary access during the setup phases was also granted to The Centre for High Performance Computing (CHPC, Cape Town, South Africa). The CHPC hardware comprises Multicore Opteron processors with InfiniBand interconnect.

2.3.4 Analysis

VMD (Humphrey *et al.*, 1996) was used for analysis of the molecular dynamics data. VMD and PYMOL (Delano, 2002) were both used for visualisation. The STRIDE program as shipped with VMD was used to assign secondary structure. The MULTISEQ plugin was used to visualise secondary structure alignments. Salt-bridges between arginine and glutamate residues were measured between the arginine C_ζ and glutamate C_δ, atoms respectively. A distance of about 4 Å corresponds to the typical distance of 2.8 Å between the hydrogen bond donor (PDB naming: NH_x) and acceptor atoms (PDB naming: OE_x) respectively.

2.3.5 Site-directed mutagenesis (IBM)

The mutations *Pf* Arg-Glu 295 Ala, *Pf* Arg-Glu 295 Arg, *Pf* Arg-Arg 404 Ala as well as the double mutation *Pf* Arg-Glu 259 Ala/Arg 404 Ala were introduced into the recombinant *P. falciparum* arginase by site-directed mutagenesis; activity and oligomeric status were detected as described in Müller *et al.* (2005). This work was performed by the collaborating group of Prof. Walter at the Biochemistry department of the Berhnard-Nocht Institute for Tropical Medicine, Hamburg, Germany.

2.3.6 Simulation of mutants

SCWRL3 (Canutescu *et al.*, 2003) was used to introduce mutations into the homology model prior to addition of hydrogens and molecular dynamics. The same protocol was then followed as for the NPT 2 simulations.

2.4 Results and discussion

2.4.1 Sequence alignment and homology modeling

2.4.1.1 *Plasmodium*-specific inserts

Searching the online *Plasmodium* genome resource, PlasmoDB (Bahl *et al.*, 2003), revealed the arginase sequences for *P. vivax*, *P. yoelii*, *P. knowlesi* and *P. berghei*, in addition to the previously characterised *PfArg* (Müller *et al.*, 2005). From the automated alignments, two *Plasmodium*-specific inserts were identified. Alignments generated from CLUSTALW and FUGUE were used as references during manual adjustment of the alignments used for modeling. In both reference cases, the positions of both inserts do not differ markedly. According to the alignment ultimately used for modeling, the first insert runs from residues 77 to 151 (75 residues), and the second from residues 377 to 388 (12 residues). Insert 1 varies considerably in sequence and length between different *Plasmodium* species, ranging from approximately 100 residues in *P. vivax* to only 15 residues in *P. berghei*. It is predicted to lie between the second β-strand and second α-helix on the outer edge of the trimer (Fig. 2.8). In contrast, insert 2 is highly conserved in sequence and length in all *Plasmodium* species. Insert 2 is located between the last β-strand and the last α-helix. The sequence identity between the *P. falciparum* and templates was 35%, 30% and 27%, respectively for the bacterial, rat and human templates. A shift of one residue for insert 2 had significant effects on its conformation relative to the trimer. When modeling according to CLUSTALW based alignments, insert 2 was predicted to interact at the trimer interface. However, when using FUGUE based alignments insert 2 was predicted to fold back away from the trimer interface and interact with it's respective monomer (Fig. 2.8).

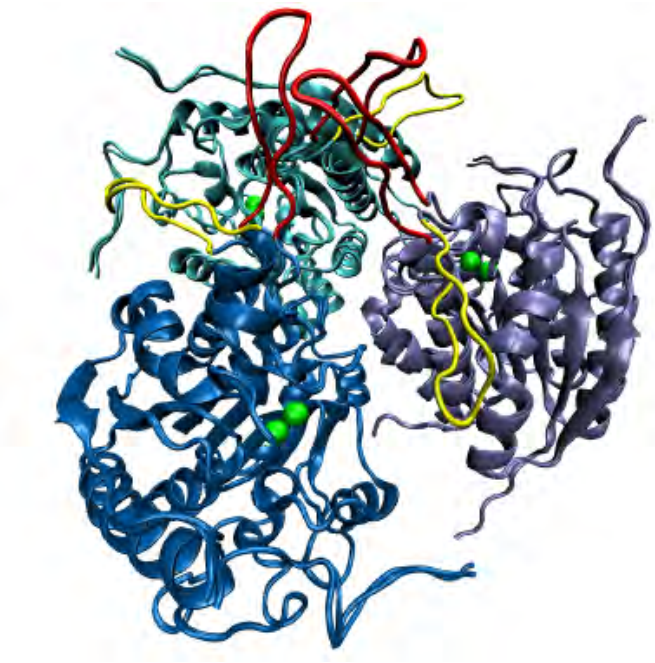


Figure 2.8: Effect of alignment on conformation of insert 2. When moving one residue upstream, the insert folds away from the trimer interface (yellow) compared to making contact (red). The active site Mg^{2+} atoms are indicated in green. Monomers are distinguished by different shades of blue.

Proteins from *Plasmodium* frequently have large inserts relative to sequences from homologues in other organisms (Aravind *et al.*, 2003). They are often characterised by low com-

plexity (Pizzi and Frontali, 2001; Xue and Forsdyke, 2003) and/or have a strong amino acid bias towards small and hydrophilic residues. The insert-originating mechanism remains unknown. Possible global functions of these inserts include nucleic acid level adaptations (Pizzi and Frontali, 2001; Xue and Forsdyke, 2003), immune evasion and protein-protein interactions (Aravind *et al.*, 2003), however these remain speculative. It has been demonstrated that interfering with these inserts can affect the protein, and thus they may have local functions relative to their enzymes (Yuvaniyama *et al.*, 2003; Birkholtz *et al.*, 2004; Jean *et al.*, 2005; Rattanachuen *et al.*, 2009). These inserts tend to align with non-core regions in homologues for which the structure is known. Most of insert 1 was left un-modelled (residues 81 to 147 removed) due to its length, leaving a small overhang at each end with respect to the templates used (Fig. 2.6-2.7). Insert 2, however, being much shorter was retained for *ab initio* modeling. A large number of sequences were used to generate multiple alignments used as references during the editing of alignments for modeling. While the reference alignments were often highly redundant, all sequences were retained in order to offset the presence of five *Plasmodium* sequences. These were included to aid in delineation of *Plasmodium*-specific inserts.

Because a small change in alignment had a substantial effect on insert 2, it is important to justify the choice of reference alignment. The FUGUE program makes use of environment-specific substitution tables and structure-dependent gap-penalties, and is thus generally expected to give a more accurate starting alignment for modeling purposes. Furthermore, the CLUSTALW alignment was also found to disturb important inter-monomer interactions. Preference was therefore given to the FUGUE reference when adjusting alignments used for modeling.

2.4.1.2 Active site and inter-monomer residue conservation

The model preserves standard active site residues observed in other arginase structures. All Mn^{2+} co-ordinating residues previously identified are conserved in the model (Section 2.4.3.6). The only substitutions are in second shell ligands when compared with the bacterial template, where Ser 176 and Glu 268 (*B. caldevelox*) are replaced by Asp 272 and Asp 365 (*P. falciparum*), respectively. Residues implicated in substrate binding are also highly conserved. There are no substitutions compared to the bacterial template and only one compared to the mammalian templates: Thr 135 (rat) is replaced by Ser 227 in the model. The high conservation of the active site suggests that inhibitors specific to the *P. falciparum* active site will be difficult to find. Thus, an alternative means of inhibition may be necessary if *P. falciparum* arginase is to be of potential therapeutic value. Therefore, attention was directed at the inter-monomer interactions.

2.4.1.3 Inter-monomer interactions

The main interaction for oligomerisation of arginase into the trimer was found to be a salt-bridge between Arg 255 and Glu 256 in adjacent monomers of human arginase I (Sabio *et al.*,

2001). This interaction is conserved in all arginase structures studied so far: Arg 249$_a$::Glu 250$_b$ (*B. caldevelox*), Arg 255$_a$::Glu 256$_b$ (human and rat arginase I), Arg 274$_a$::Glu 275$_b$ (human and rat arginase II). In *P. falciparum,* the corresponding interaction is formed by Arg 346$_a$::Glu 347$_b$.

Visual inspection of the models suggested a novel inter-monomer salt-bridge forms between Glu 295$_a$::Arg 404$_b$ (Fig. 2.9). In multiple sequence alignments, Glu 295 aligns with conserved acidic residues in the bacterial and mammalian templates. *P. falciparum* Glu 295 aligns with an Asp in mammals (human arginase II: Asp 223, rat arginase I: Asp 204), fungi and bacteria (*B. caldovelox* arginase: Asp 199). In the other *Plasmodium* species, Glu 295 aligns either with Asp (*P. yoelii* and *P. berghei*) or Glu (*P. knowlesi* and *P. vivax*). The only exception is in plants, where a serine residue is found instead (*Arabidopsis thaliana* arginase I: Ser 247). In the model, and the mammalian and bacterial templates this acidic residue forms interactions with the adjacent monomer via partner residues that do not align. In mammals, Asp 223/204$_a$ forms a salt bridge with Arg 308/327$_b$ (rat arginase I/human arginase II) instead. This salt-bridge nucleates considerable inter-monomer interactions, characterised by an S-shaped *C*-terminus (Kanyo *et al.*, 1996; Cama *et al.*, 2003a). The *Plasmodium* sequences lack such an extensive *C*-terminus and Arg 308 aligns with a hydrophobic Ile instead (408: *P. falciparum*, 368: *P. knowlesi*, 436: *P. vivax*, 353: *P. berghei*, 376: *P. yoelii*). In the bacterial structure the Asp 199 cognate forms an inter-monomer bridge with Glu 256 that is mediated either by urea or by free arginine, depending on the crystallisation conditions (Bewley *et al.*, 1999). The *Plasmodium* Arg salt-bridge partner to Glu 295 aligns with small and/or hydrophilic residues in other organisms (eg. Ser, Thr, Cys, Ala, Glu). The importance of the S-shaped tail is still in doubt since products truncated after Arg 308 can still form active trimers (Mora *et al.*, 2000). In bacterial structures, an interaction is formed with another acidic residue (Glu 256) that is mediated by either urea or free arginine. Finally, in the *P. falciparum* model Glu 295 is predicted to interact with Arg 404. Arg 404 doesn't align with Arg 308 or Glu 256. Thus, there appears to have been evolutionary pressure to establish a strong inter-monomer interaction in this region of the monomer-monomer interface. The differences between the *P. falciparum* model and templates suggest this salt bridge as a possibly unique interaction and was therefore subject to scrutiny using molecular dynamics and site-directed mutagenesis.

In order to determine other possible interactions the salt bridge analysis tool of VMD was used to search for all possible salt bridges in the protein, using co-ordinates prior to sampling. All salt bridges with a hydrogen-bond donor/acceptor distance less than 3.2 Å were identified. Only one other interaction between adjacent monomers was found, between Glu 400 and Lys 340. However, this interaction was not stable during molecular dynamics. This was observed both with and without Mg^{2+} (results not shown). Thus, this interaction is likely to be only of secondary importance in maintaining quaternary structure, and attention was focused on the Glu 295::Arg 404 interaction instead.

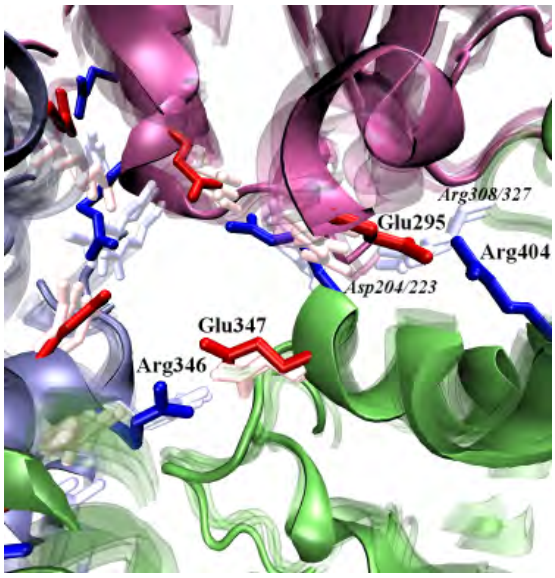


Figure 2.9: Inter-monomer salt bridges in *Pf*Arg. The conserved inter-monomer salt-bridge at Arg 346$_a$::Glu 347$_b$ and the non-conserved inter-monomer salt-bridge at Glu 295$_a$::Arg 404$_b$ are both indicated. Monomers are in green, light-blue and mauve. The rat arginase 1 and human arginase II templates are shown in transparency, and their respective salt-bridges indicated in lighter shades of red (acidic) and blue (basic). Template residues are in italics (*rat/human*). An interactive 3D summary of the various inter-monomer interactions is given in Fig. A.1

2.4.2 Initial simulations

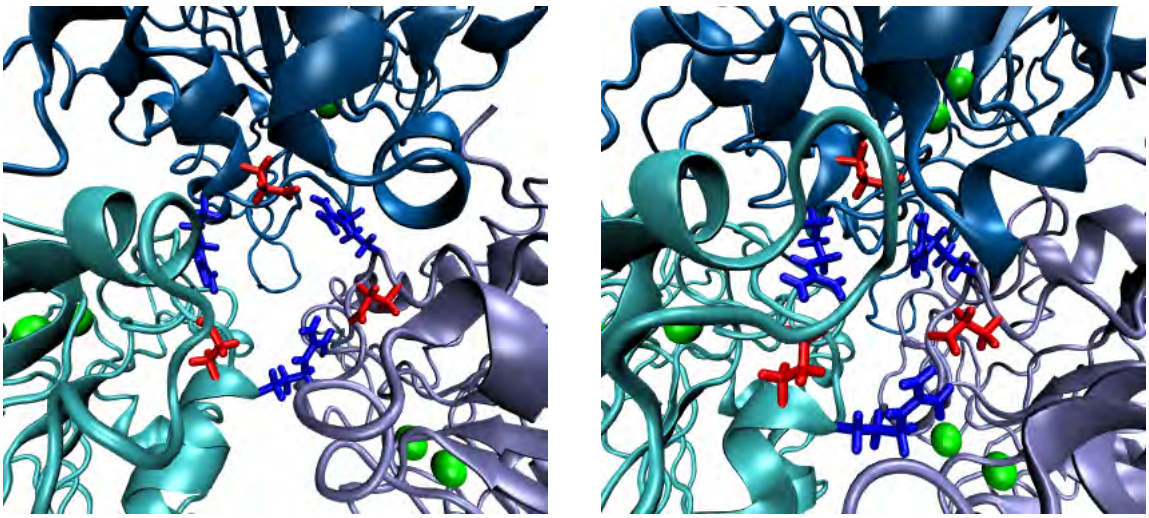
2.4.2.1 Simulation times with CHARMM

Simulating the complete trimer with CHARMM was found to be quite slow, even in parallel. Using 4 CPUs, heating the shell solvated system (2.3.3.1) took approximately 14 hours for 30 000 steps (± 55 000 atoms). This amounts to 1.6 s of wall-clock time per 1 fs timestep in the simulation. The equivalent simulation of the trimer without solvent (± 16800 atoms 2.3.3.1, 2.3.3.1) took approximately 1.35 hours with 16 CPUs (0.18 seconds per timestep (1 fs)), 1.45 hours with 12 CPUs (0.3 seconds per timestep (1 fs)) and 2.5 hours with 4 CPUs (1.44 seconds per timestep (1 fs)). Simulating models based on the a2 alignment (± 15900 atoms) on 8 CPUs gave a time of about 0.23 s per timestep (1 fs). By comparison, simulating ± 96 000 atoms with NAMD using 4 CPUs took 1.4 s per timestep (2 fs). Thus, with CHARMM simulating comparatively small systems took a relatively long time. While the shell solvated system is relatively small, it will not remain intact. The solvation shell attempt to lower surface tension over time and this converge on a spherical droplet modified by the presence of the protein solute. Furthermore, individual solvent molecules will drift away from the system due to random chance. Thus, it is not suitable for accurate simulations with explicit solvent. Instead, explicit solvation requires appropriate boundary conditions that will ensure the protein surface is in contact with solvent. Because of the slow simulation times that would have accompanied a fully solvated protein box, explicit solvation was abandoned for simulations in CHARMM.

2.4.2.2 Trimer interface integrity with CHARMM

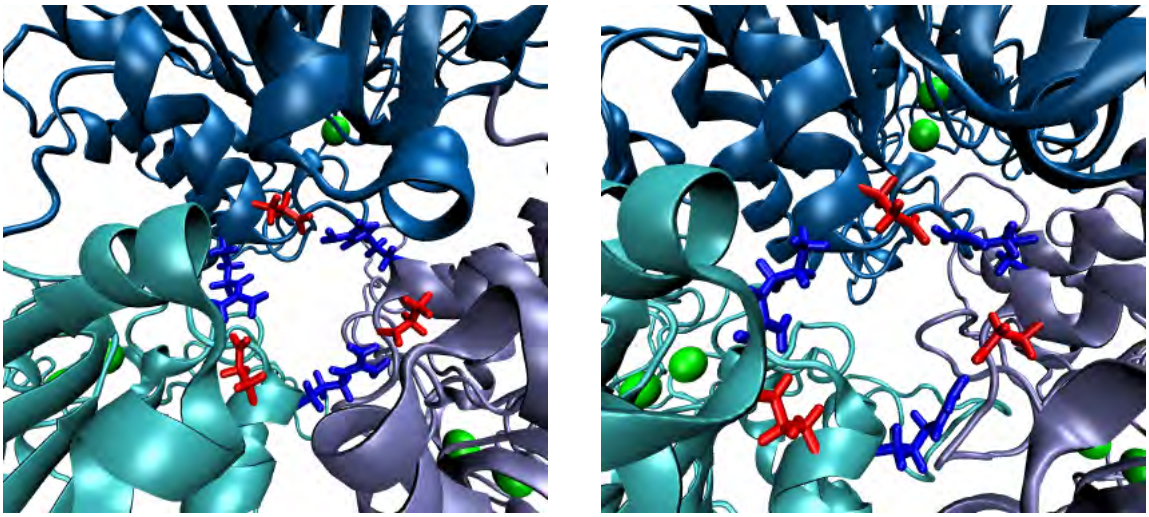
The salt-bridge at Arg 346$_a$::Glu 347$_b$ is conserved in all known arginase structures. It is required for maintaining the trimer complex, and hence it's integrity during MD simulations was closely monitored. During *in vacuo* equilibration of the early model this interaction was disturbed or poorly maintained (Fig. 2.10). The trimer interface was also observed

to collapse during *in vacuo* simulations. Imposing symmetry on the internal co-ordinates was observed to improve this slightly, but there was still collapse of the trimer interface (Fig. 2.11). Using models built on the a2 alignment without symmetry did not improve the situation at the trimer interface either (Fig. 2.12). Combining both the use of the a2 alignment and symmetry was observed to have less of a detrimental effect on the trimer interface, even after 90 000 steps (90 ps) for systems both with and without Mg^{2+}. Therefore, these simulations were extended for production/sampling runs. However, after 10 ns the same kind of collapse at the trimer interface was observed (Fig. 2.13).



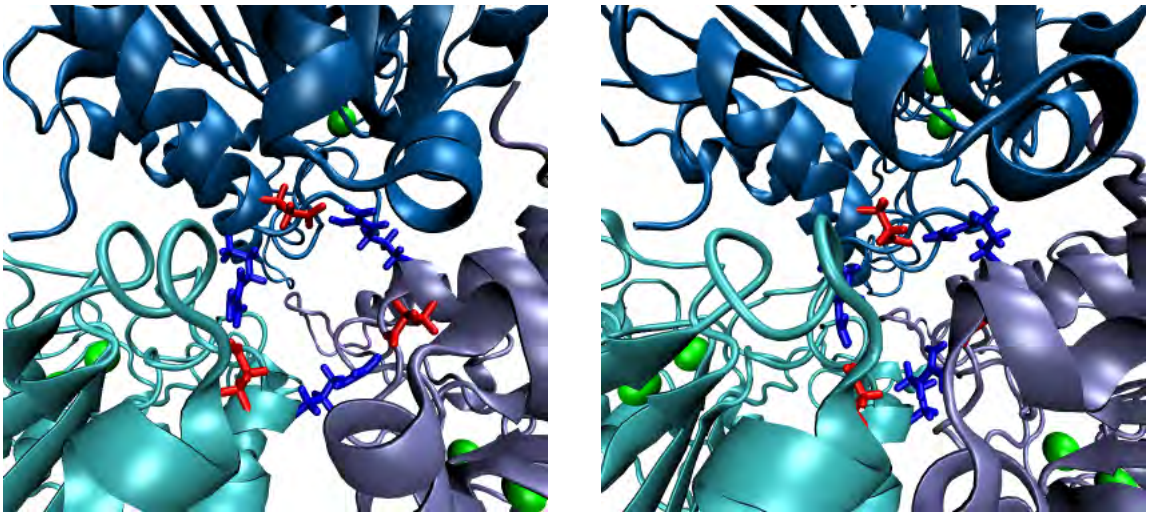
(a) Before equilibration (b) After equilibration

Figure 2.10: *In vacuo* equilibration in CHARMM with an early model. During equilibration the trimer interface usually collapses, which disturbs the conserved interaction at Arg 346_a::Glu 347_b Highlighted residues (blue: Arg, red: Glu) are shown as stick representations. The different chains are indicated in different shades of blue. The active site metals are depicted as green spheres.



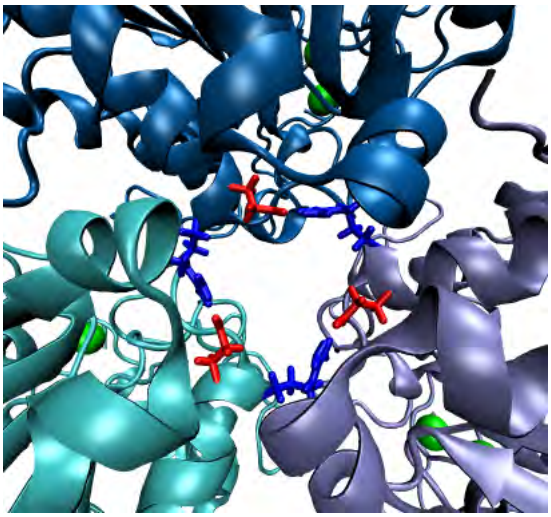
(a) Before equilibration (b) After equilibration

Figure 2.11: *In vacuo* equilibration in CHARMM with an alignment 1 (a1) derived model including symmetry. During equilibration the trimer interface is still observed to collapse, however less so than with the MD on the early model. The conserved interaction between Arg 346_a::Glu 347_b is indicated. Highlighted residues (blue: Arg, red: Glu) are shown as stick representations. The different chains are indicated in different shades of blue. The active site metals are depicted as green spheres. The active site metals are depicted as green spheres.

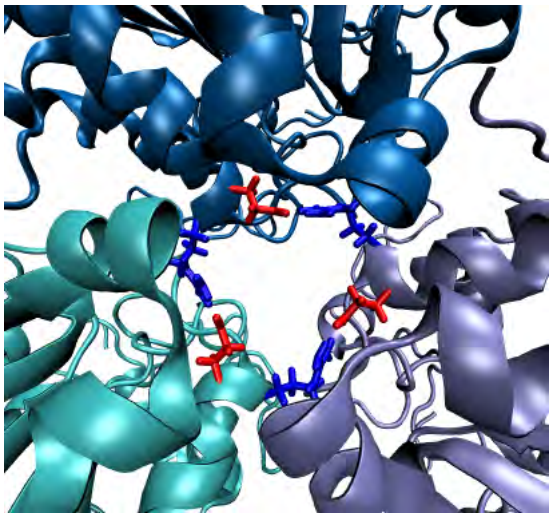


(a) Before equilibration (b) After equilibration

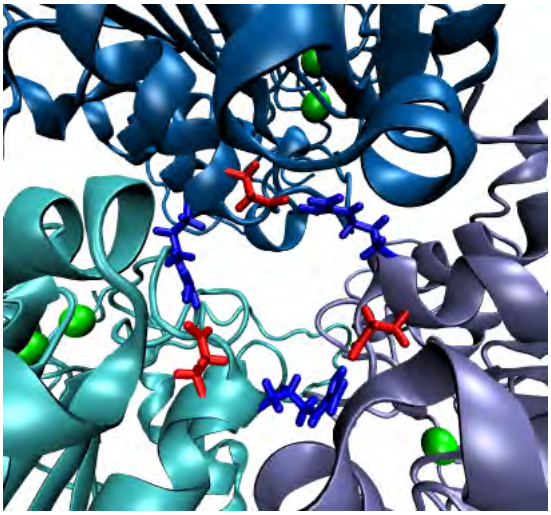
Figure 2.12: *In vacuo* equilibration in CHARMM with alignment 2 (a2) derived model excluding symmetry. During equilibration the trimer interface collapse is again observed. The conserved interaction between Arg 346_a::Glu 347_b is indicated. Highlighted residues (blue: Arg, red: Glu) are shown as stick representations. The different chains are indicated in different shades of blue. The active site metals are depicted as green spheres.



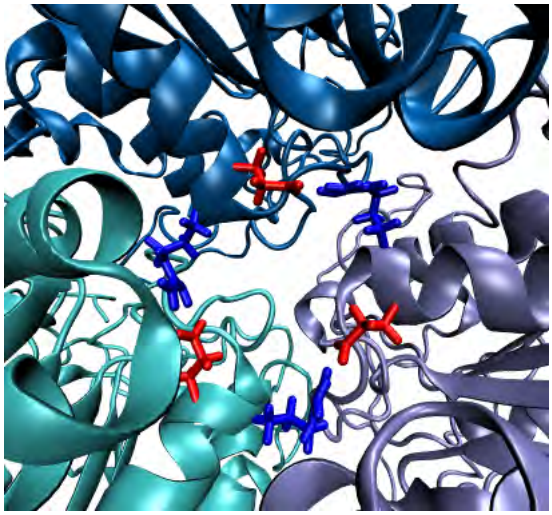
(a) Before equilibration (+ Mg^{2+})



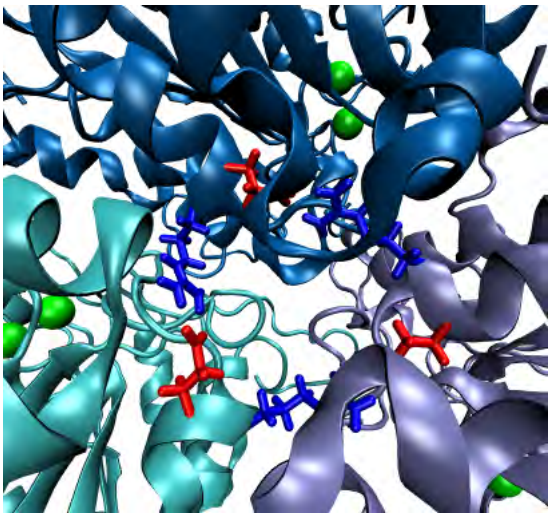
(b) Before equilibration (- Mg^{2+})



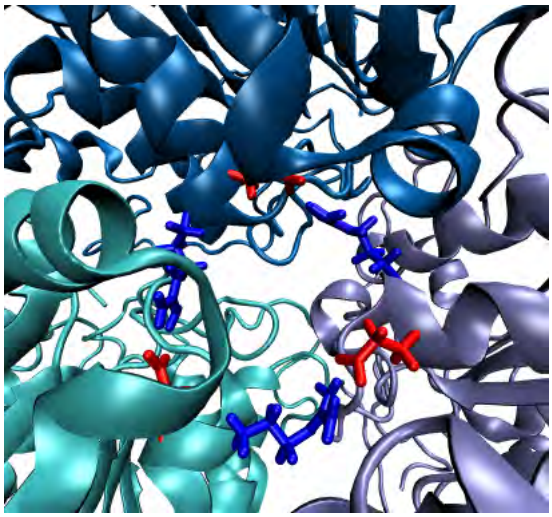
(c) After equilibration (+ Mg^{2+})



(d) After equilibration (- Mg^{2+})



(e) After 10 ns sampling (+ Mg^{2+})



(f) After 10 ns sampling (- Mg^{2+})

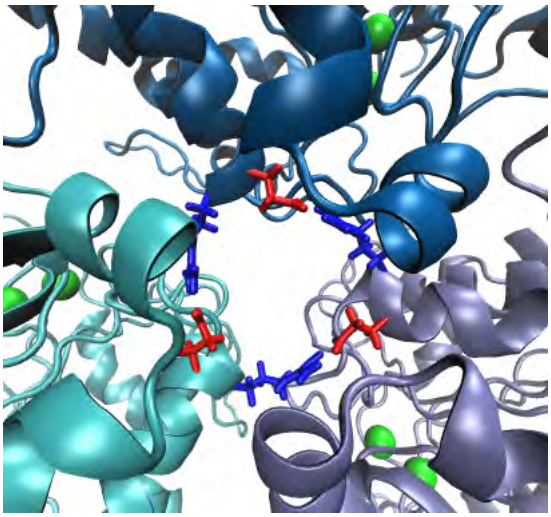
Figure 2.13: *In vacuo* equilibration with alignment 2 (a2) derived including symmetry. During equilibration the trimer interface collapse is again observed and worsens during sampling. The conserved interaction at Arg 346_a::Glu 347_b is indicated. Highlighted residues (blue: Arg, red: Glu) are shown as stick representations. The different chains are indicated in different shades of blue. The active site metals are depicted as green spheres.

2.4.2.3 *In vacuo* simulations with NAMD

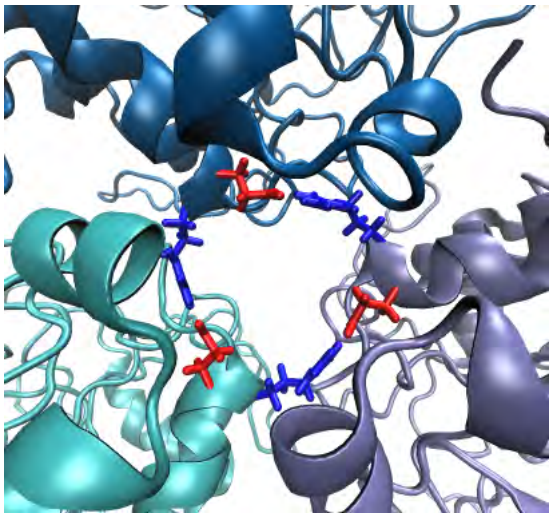
Similar behaviour was observed as for the CHARMM simulations. At the end of equilibration one of the Arg 346_a::Glu 347_b salt-bridges was broken in the simulation with Mg^{2+}, while all three salt-bridges were still intact in the system without Mg^{2+}. Collapse of the central pore was also observed in both systems. After 1.5 ns of the production run this was more pronounced in both cases and all three salt-bridges were broken. There was also more visible distortion and loss of secondary structure in the simulation without Mg^{2+} (Fig. 2.14). By the end of 12 ns of the production run (executed for the non-Mg^{2+}simulation only) the destabilisation of the protein had increased further.

Although both systems with and without Mg^{2+} showed considerable distortion during molecular dynamics, there was more distortion in the non-metal system. This is reflected in the Root Mean Square Deviation (RMSD) of the positions of the $C_\alpha$atoms during the production run. In the system with Mg^{2+}removed, the RMSD compared to the first frame shows a greater increase (Fig. 2.15). This indicates that there is more structural distortion of the trimer complex with the active site metal removed.

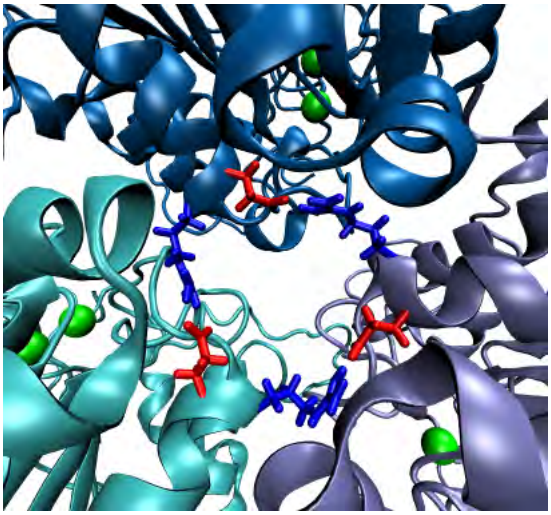
Simulating *in vacuo* with NAMD was judged to be inadequate for investigating the structural metal dependency. Firstly, because the conserved interaction at Arg 346_a::Glu 347_b was not stable in the simulation with Mg^{2+}, and secondly because of the collapse of the central pore of the trimer interface after only a short time of simulation. These results further confirmed the necessity of having solvent in the central pore to prevent collapse.



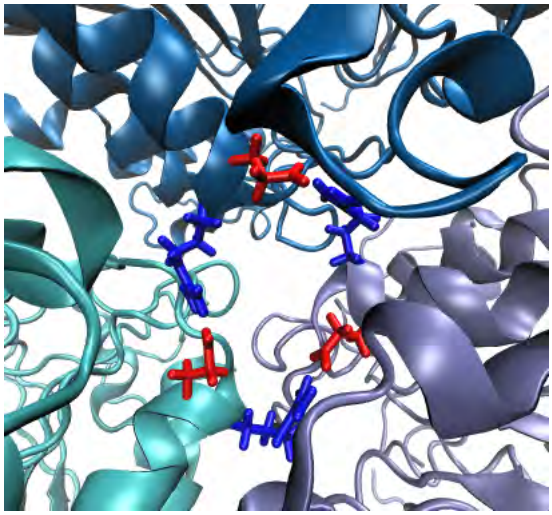
(a) Before equilibration (+ Mg^{2+})



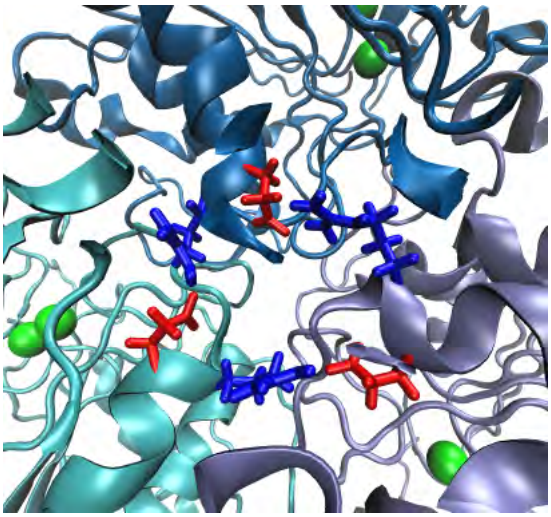
(b) Before equilibration (- Mg^{2+})



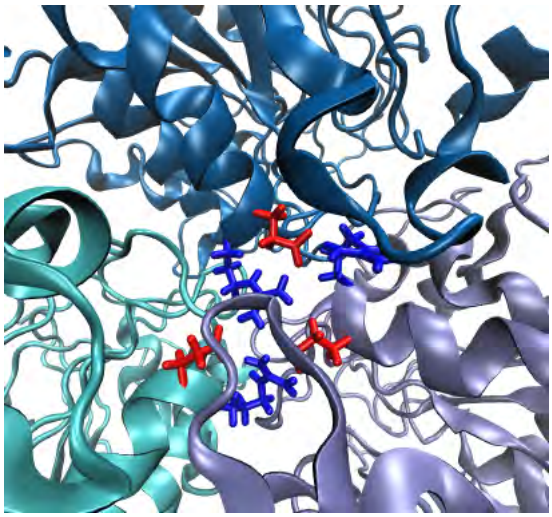
(c) After equilibration (+ Mg^{2+})



(d) After equilibration (- Mg^{2+})



(e) After 1.5 ns sampling (+ Mg^{2+})



(f) After 1.5 ns sampling (- Mg^{2+})

Figure 2.14: *In vacuo* simulations with NAMD - a2 alignment (with symmetry and metals). During equilibration the trimer interface collapse is again observed. The conserved interaction at Arg 346_a::Glu 347_b is indicated. Highlighted residues (blue: Arg, red: Glu) are shown as stick representations. The different chains are indicated in different shades of blue. The active site metals are depicted as green spheres.

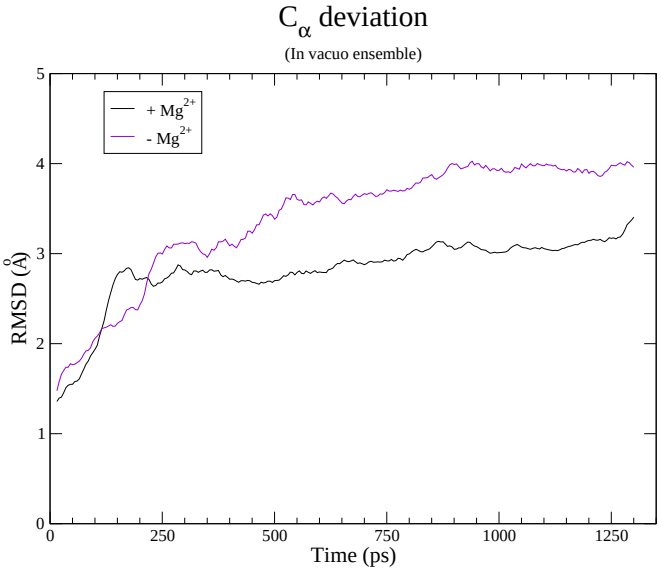


Figure 2.15: C_α RMSD during the production run of the *in vacuo* ensemble with and without Mg^{2+}. Removing Mg^{2+} causes a greater increase in RMSD as sampling progresses. The plots were averaged using a sliding window of 50 frames (500 fs).

2.4.2.4 Conclusions

In general it was found that *in vacuo* simulations in CHARMM were not suitable for investigating the structural metal dependency. In all simulations the trimer interface was observed to collapse, often destroying the integrity of the Arg 346_a::Glu 347_b salt-bridge which is required for trimer integrity. For reliable simulations this interaction should remain intact in conditions not expected to disturb the trimer. The interface collapse is likely due to the absence of any solvent species occupying the centre of the trimer which would help prevent collapse through steric interactions. The presence of explicit solvent would also shield strong non-bonded interactions across the central pore. Furthermore, the CHARMM simulations were found to be fairly slow, with poor scaling when running on larger numbers of CPUs. During this, NAMD was also tested which is known to scale well on the kind of hardware that was available. The better performance of NAMD is probably due to the optimised message-passing and spatial decomposition algorithms for commodity-hardware based clusters (Phillips *et al.*, 2005). It was found that fully solvated NAMD simulations (± 96 000 atoms), executed at approximately the same speed as un-solvated CHARMM simulations (± 16 000 atoms). However, as with CHARMM, the integrity of the trimer interface during *in vacuo* simulations was also not maintained with NAMD. Therefore, because the execution speed was the same and including solvent is more biologically realistic, it was decided to do all further simulations with NAMD and explicit solvent.

2.4.3 Solvated simulations

2.4.3.1 Introduction of chain break for insert1

Before proceeding with detailed analysis, it was necessary to ensure that gross changes to the protein necessitated by the modeling difficulties would not compromise the interpretation of results. The omission of insert 1 is potentially problematic, in that it introduces an unnatural

chain-break and therefore potential instability into the protein. The deletion of the insert creates a protein fragment on the outer edge of the trimer complex that does not interact extensively with any neighbouring monomers, and makes most of its contacts with it's own monomer. The insert was too long to attempt any modeling of this using the protocols described. The gap was left un-ligated, and apart from the loss of some secondary structures in this region (described below), this fragment was stable for up to 50 ns of simulation and remained in contact with the rest of the protein. The deletion of insert 1 for modeling did not adversely affect the stability of the model: while potential problems by introducing a chain break could have been avoided by ligating the ends of the gap, this would also be unnatural. Because the fragment was apparently stable and closing the gap un-ligated is less parsimonious, the break was left in.

2.4.3.2 RMSD

Protein stability was monitored by the change in C_α RMSD during the sampling runs. C_α RMSD was compared with the first frame during the equilibration and sampling runs. In all solvated simulations there is an increase in RMSD which equilibrates after about 10 ns for the NP simulation, and between 20 and 30 ns for the NPT simulations. In two out of three (NP: Fig. 2.16 , NPT_2: Fig: 2.18), the C_α RMSD equilibrates at 1 Å more than in the system with Mg^{2+}, similar to what was observed for the *in vacuo* simulation. In the NPT_1 simulation the situation is more variable, with the RMSD of -Mg^{2+} less than that of the Mg^{2+} system between 10 and 30 ns, but larger for the first 10 and last 5 ns out of 55 ns of simulation (Fig. 2.17). The equilibration of C_α RMSD at a larger distance for the -Mg^{2+} systems suggests that removing the active site metals has a detrimental effect on protein stability. It was previously reported that removing Mn^{2+}, either by dialysis and chelation with EDTA, or by mutagenesis of Mn^{2+} co-ordinating residues in the active site, not only abolished enzyme activity but also promoted dissociation of the trimer into monomers. Reformation of the trimer could be obtained by addition of Mn^{2+} (Müller *et al.*, 2005).

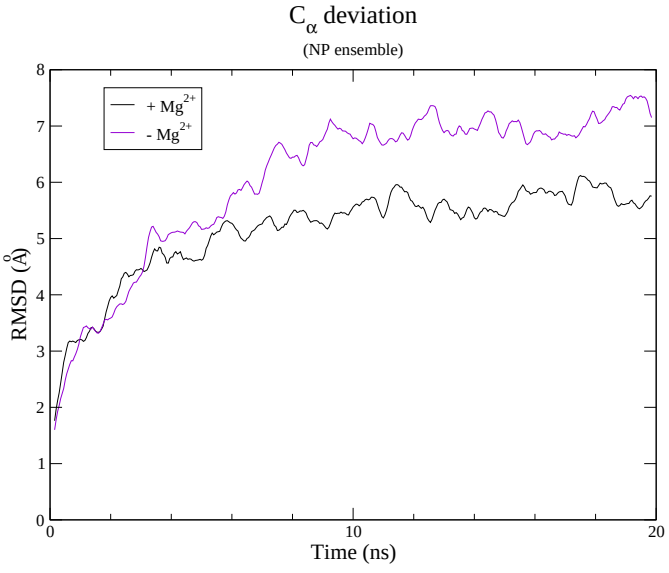


Figure 2.16: C_α RMSD during the production run of the NP with (+) and without (-) Mg^{2+}. Removing Mg^{2+} causes a greater increase in RMSD as sampling progresses. The plots were averaged using a sliding window of 500 frames (250 fs).

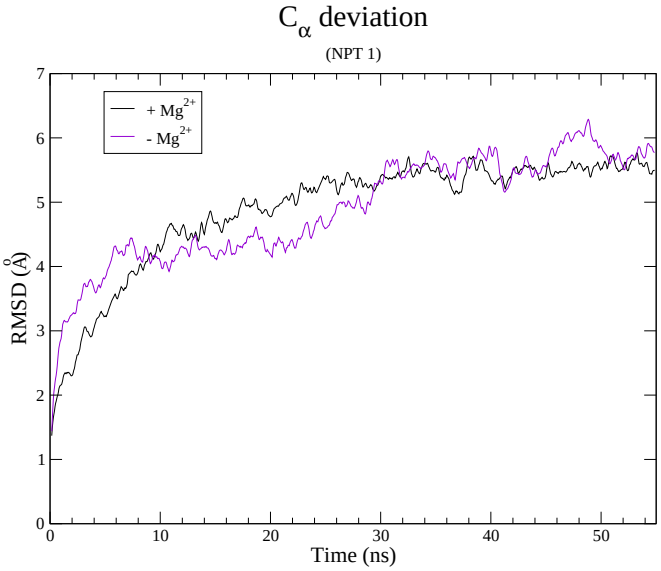


Figure 2.17: C_α RMSD during the production run of the NPT_1 with (+) and without (-) Mg^{2+}. There was no noticable difference with and without Mg^{2+} . The plots were averaged using a sliding window of 500 frames (250 fs).

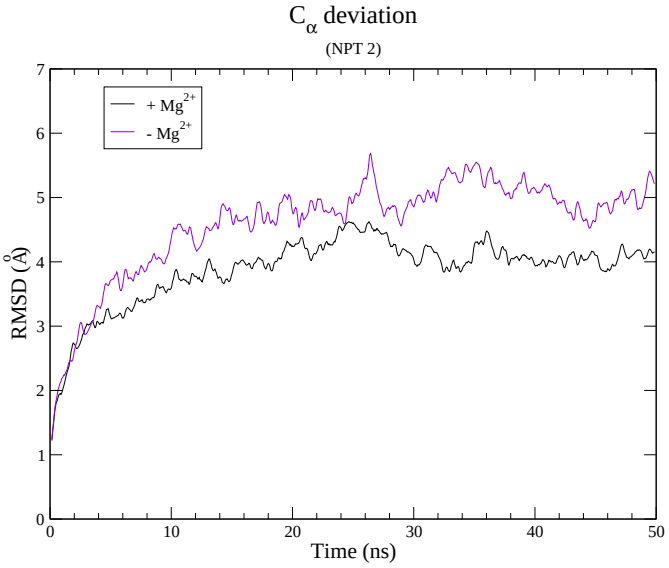


Figure 2.18: C_α RMSD during the production run of the NPT_2 with (+) and without (-) Mg^{2+}. Removing Mg^{2+} causes a greater increase in RMSD as sampling progresses. The plots were averaged using a sliding window of 500 frames (250 fs).

The RMSD of the mutant simulations was monitored as described above for the Mg^{2+} and Mg^{2+}-free simulations. In all mutants a similarly greater increase is observed compared to the wild type with Mg^{2+} (Fig. 2.19). The greatest increase occurs in the *Pf* Arg-Glu 347 Gln mutant, which equilibrates at ± 2.5 Å more than the wild type/Mg^{2+} by 40 ns. The increase is also the most rapid for *Pf* Arg-Glu 347 Gln mutant at ± 2 Å more by 10 ns. The next largest effect is observed for the Glu 295 Arg mutant, which largely mirrors the removal of Mg^{2+}, with a similar increase by 6 ns as for the Glu 347 Gln mutant. By 40 ns the RMSD for *Pf* Arg-Glu 295 Arg is 0.8 Å larger than wild type/Mg^{2+}. The *Pf* Arg-Glu 295 Ala, *Pf* Arg-Arg 404 Ala and *Pf* Arg-Glu 295 Ala/Arg 404 Ala mutants display a similar increase in RMSD by 40 ns, between 0.5-0.8 Å larger than for Mg^{2+}.

2.4.3.3 Preservation of secondary structure

The effect of removing the metal on conservation of secondary structure (Fig. 2.20) during sampling was also monitored. In general, a greater loss of secondary structural integrity was observed for non-metal systems in the NP and NPT_1 and NPT_2 ensembles (Table 2.3).

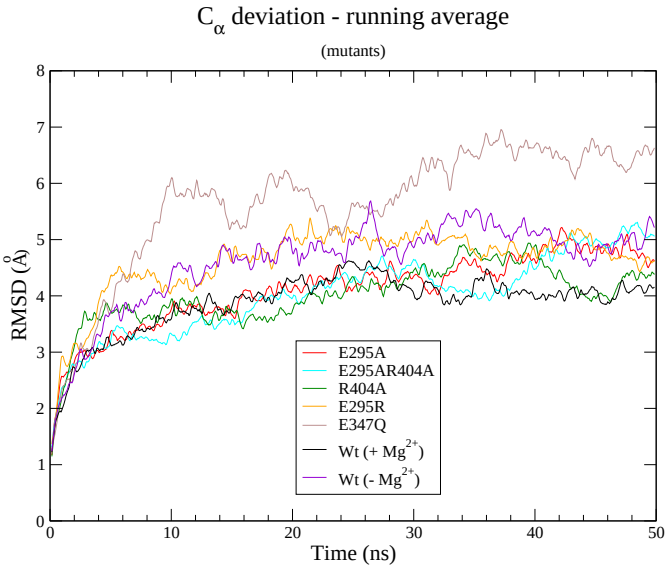


Figure 2.19: C_α RMSD of E295A, E295R, E347Q, R404A and E295AR404A mutants compared to wild type with Mg^{2+}. Introducing mutations into the inter-monomer salt-bridges causes a similar increases in equilibrium RMSD compared to removing Mg^{2+}. The plots were averaged using a sliding window of 500 frames (250 fs).

Combining all the simulations gives a total of 9 simulations for the monomer in each condition ($\pm Mg^{2+}$), which can be used to observe any general loss of secondary structure. However, these data are not sufficient to determine possible co-operative effects between monomers. In the absence of Mg^{2+}, complete loss of larger secondary structure is observed for few elements. The monomers/chains are arbitrarily designated A, B and C. In chain B of the NP simulation and chain C of NPT_1, the second β-strand is lost. The second α-helix of chain C in NPT_2 is also lost. For some of the smaller secondary structure elements ($\leq$4 residues) there is complete loss of structure for both conditions. This is observed for $\beta 3$, $\beta 4$, $\alpha 6$, $\alpha 10$ and $\alpha 12$. All of these secondary structural elements align with similar elements in all of the templates. The greater loss of secondary structure without $Mg^{2+}$is expected from the greater increase in RMSD. However, from this data there is no obvious protein region that suggests further investigation with regards to disturbing arginase activity.

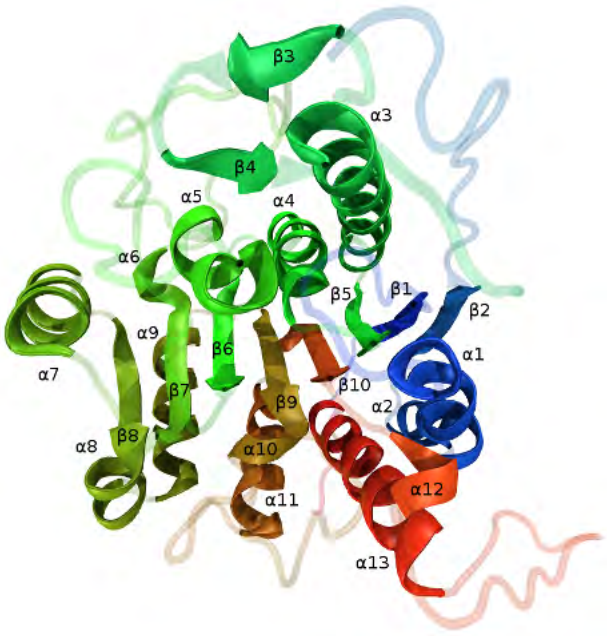


Figure 2.20: Tertiary structure of the *Plasmodium* arginase model. Secondary structure elements (αx for strands, βx for helices) are indicated. Increasing sequence position is indicated by a bluegreenred colour gradient. Loops and coiled regions are transparent for clarity.

Table 2.3: Preservation (+) and loss (-) of secondary structure during molecular dynamics. Each string approximately corresponds to the length of the secondary structure element (β-strands: 3-4 residues, $\alpha/3_{10}$-helices: 4-5 residues). The sum total of symbols is 450. A greater loss of secondary structure is observed upon removing Mg^{2+}: lost structure in simulations is represented by 48 and 65 "-" symbols in $+Mg^{2+}$and $-Mg^{2+}$simulations, respectively.

Condition	Simulation	Chain	$\beta1$	$\alpha1$	$\alpha2$	$\beta2$	$\beta3$	$\beta4$	$\alpha3$	$\beta5$	$\alpha4$	$\beta6$	$\alpha5$	$\alpha6$
$+Mg^{2+}$	NP	A	++	+++	-++	++	-	-	++++	++	+++	++	-+	-
		B	+-	+++	+++	+-	-	-	-+++	++	+++	++	-+	+
		C	++	+++	+++	++	-	-	++++	++	+++	++	-+	-
	NPT_1	A	++	-++	+++	++	-	-	++++	++	+++	++	-+	+
		B	++	+++	+++	++	-	-	-+++	++	+++	++	-+	+
		C	+-	-++	+++	+-	-	-	++++	++	+++	++	-+	-
	NPT_2	A	+-	+++	++-	+-	-	-	++++	++	+++	++	++	+
		B	+-	-++	+++	+-	-	-	++++	+-	+++	++	-+	-
		C	+-	+++	+++	+-	+	+	++++	++	+++	++	-+	+
		$\frac{-}{total}$	$\frac{5}{18}$	$\frac{3}{17}$	$\frac{2}{27}$	$\frac{5}{18}$	$\frac{8}{9}$	$\frac{8}{9}$	$\frac{2}{36}$	$\frac{1}{18}$	$\frac{0}{27}$	$\frac{0}{18}$	$\frac{8}{18}$	$\frac{4}{9}$
$-Mg^{2+}$	NP	A	+-	+++	+++	+-	-	-	++++	++	+++	++	++	+
		B	+-	+++	+++	--	-	-	--++	++	+++	++	++	-
		C	++	+++	+++	++	-	-	+-++	+-	+++	++	-+	-
	NPT_1	A	+-	-++	-++	+-	-	-	++++	++	+++	++	++	+
		B	+-	-++	+++	+-	-	-	++++	++	+++	++	++	+
		C	+-	+++	+++	--	-	-	-+++	++	+++	++	-+	-
	NPT_2	A	+-	+++	+++	+-	-	-	--++	+-	+++	++	-+	+
		B	+-	+++	+++	+-	-	-	++++	+-	+++	++	-+	+
		C	+-	+++	---	+-	+	+	++++	++	+++	++	-+	+
		$\frac{-}{total}$	$\frac{8}{18}$	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{10}{18}$	$\frac{8}{9}$	$\frac{8}{9}$	$\frac{6}{36}$	$\frac{3}{18}$	$\frac{0}{27}$	$\frac{0}{18}$	$\frac{5}{18}$	$\frac{3}{9}$

Table 2.3: Table 2.3 continued...

Condition	Simulation	Chain	β7	α7	β8	α8	α9	β9	α10	α11	β10	α12	α13
$+Mg^{2+}$	NP	A	++	++	++	++	+++	++	-	+++	++	-	-++
		B	++	++	++	++	+++	++	+	+++	++	+	-++
		C	++	++	++	++	+++	++	+	+++	++	-	-++
	NPT_1	A	+-	++	++	++	+++	++	+	+++	++	-	+++
		B	++	++	++	++	+++	++	-	+++	++	+	-++
		C	++	++	++	++	+++	++	+	+++	++	-	-++
	NPT_2	A	++	++	++	++	+++	++	+	+++	++	-	-++
		B	++	++	++	++	+++	++	+	+++	++	-	+++
		C	++	++	++	++	+++	++	+	+++	++	-	+++
		$\frac{-}{total}$	$\frac{1}{18}$	$\frac{0}{18}$	$\frac{0}{18}$	$\frac{0}{18}$	$\frac{0}{27}$	$\frac{0}{18}$	$\frac{2}{9}$	$\frac{0}{27}$	$\frac{0}{18}$	$\frac{7}{9}$	$\frac{6}{27}$
$-Mg^{2+}$	NP	A	++	++	++	++	+++	++	-	+++	-+	-	+++
		B	++	++	++	++	+++	++	+	+++	+-	-	-++
		C	++	++	++	++	+++	++	-	++-	+-	-	+++
	NPT_1	A	++	++	++	++	+++	+-	-	+++	-+	-	+++
		B	++	++	++	++	+++	++	-	+++	++	-	-++
		C	++	++	++	++	+++	++	-	+++	++	-	+++
	NPT_2	A	++	++	++	++	+++	++	+	+++	++	-	+++
		B	++	++	++	++	+++	++	-	+++	++	-	+++
		C	++	++	++	++	+++	++	+	+++	++	-	+++
		$\frac{-}{total}$	$\frac{0}{18}$	$\frac{0}{18}$	$\frac{0}{18}$	$\frac{0}{18}$	$\frac{0}{27}$	$\frac{1}{18}$	$\frac{6}{9}$	$\frac{1}{27}$	$\frac{4}{18}$	$\frac{9}{9}$	$\frac{2}{27}$

2.4.3.4 Effect of Mg^{2+}removal on movement of insert 2

During the simulation insert 2 does not retain its interaction as predicted by the original homology model prior to molecular dynamics. Considerable movement is observed instead. Furthermore, there are some noticeable differences between the metal and non-metal simulations. In general when Mg^{2+} is removed insert 2 moves towards the trimer interface between two adjacent monomers (Fig. 2.21, Fig. 2.22). In the NP $-Mg^{2+}$ simulation insert 2 on all three monomers adopts a conformation close to the monomer interfaces of the trimer. In two chains, these inserts show reasonable agreement in that they appear to occupy the interface between two monomers. In the third chain, the insert moves to the centre of the trimer instead. In the NP $+Mg^{2+}$ simulation, this is less pronounced with some movement towards the interface between neighbouring monomers only observed for one chain. The other two remain solvent exposed. In the NPT simulations, there is movement by all three inserts

to the monomer-monomer interface with Mg^{2+}, although less pronounced than for the NP simulation. In the -Mg^{2+} simulations, the inserts again remain more solvent exposed. These results suggest insert 2 may also be involved in maintaining the trimer and may be part of the structural metal dependency. Furthermore, removing the active site metals compromises global structural integrity of *P. falciparum* arginase.

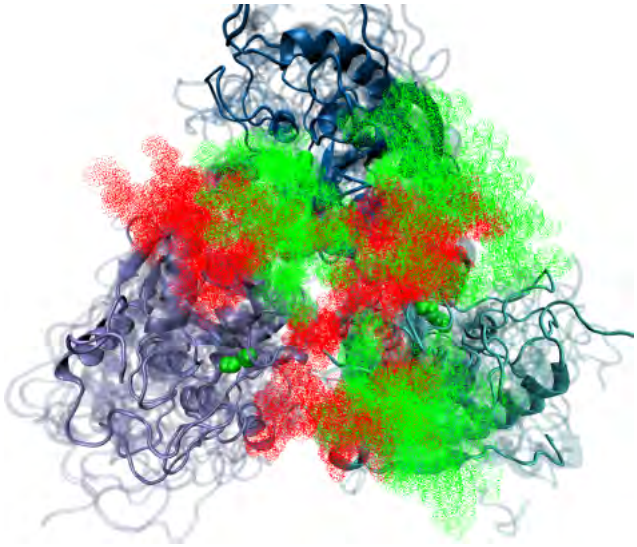


Figure 2.21: Movement of insert 2 - top view. Trimer states at the end of 20 ns (NP) and 50 ns (NPT_1 and NPT_2) are overlayed. Insert 2 is shown as clouds, Mg^{2+}(red) and non-Mg^{2+}(green). In general when Mg^{2+} is removed insert 2 moves towards the trimer interface between two adjacent monomers.

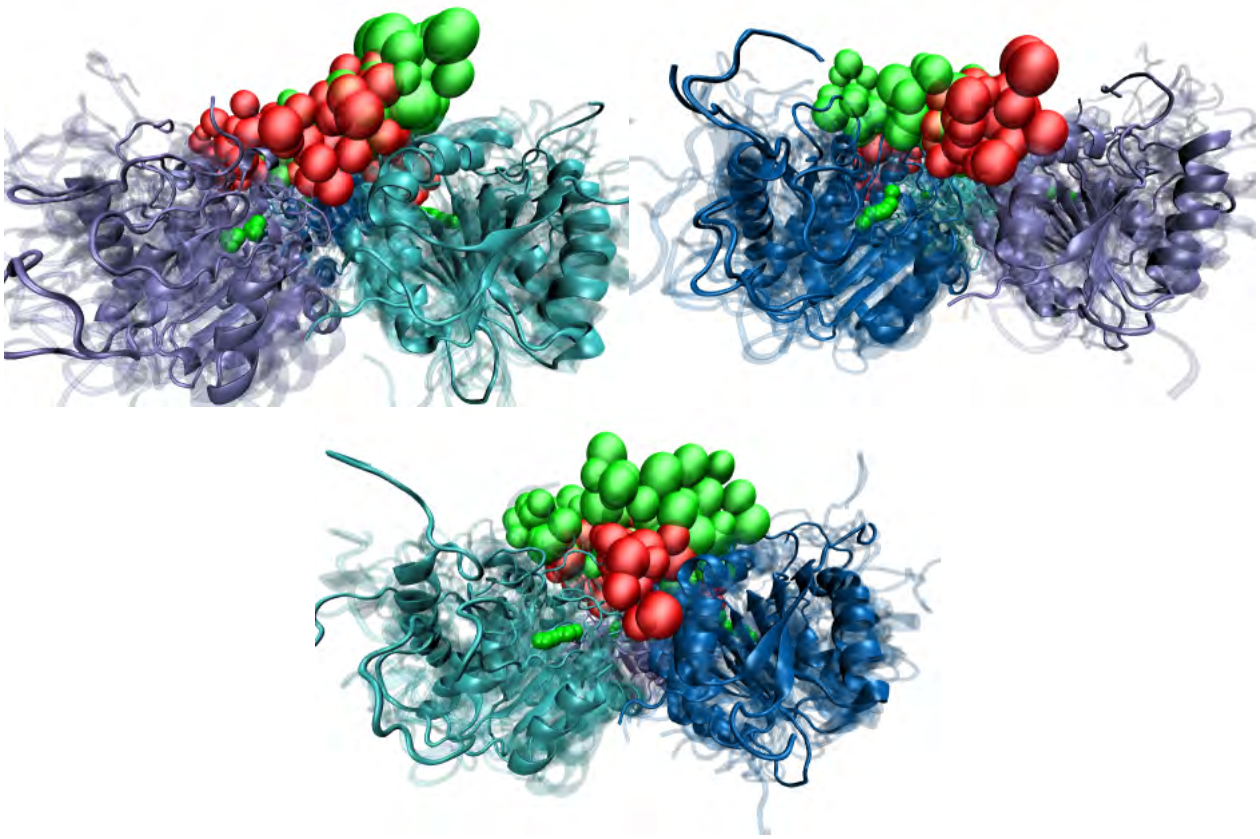


Figure 2.22: Movement of insert 2 - side on views. Trimer states at the end of 20 ns (NP) and 50 ns (NPT_1 and NPT_2) are overlayed. Insert 2 is shown as one bead per residue, +Mg^{2+}(red) and -Mg^{2+}(green). In general when $Mg^{2+}$is removed insert 2 moves towards the trimer interface between two adjacent monomers.

2.4.3.5 Integrity of inter-monomer salt-bridges

In all arginases studied to date there is a conserved inter-monomer salt-bridge represented in *P. falciparum* by Arg 346$_a$::Glu 347$_b$ (Fig. 2.9). These residues align unambiguously with their template cognates: Arg 255/274/249$_a$::Glu 256/275/250$_b$ (rat, human and bacterial templates, respectively). The salt-bridge forms reliably during modeling. Considering the established importance of this interaction, its integrity was monitored during modeling and simulation.

In the sampling runs Arg 346$_a$::Glu 347$_b$ was generally stable and intact for both the +Mg^{2+} and -Mg^{2+} case. One inter-monomer bridge did break in the Mg^{2+} case of the NP ensemble (Fig. 2.23). In the NPT ensembles the interactions mostly remained intact in both cases. One interaction was observed to break for about 6 ns in the NPT_1 -Mg^{2+} ensemble. However, there is also an increase in the average standard deviation of the salt-bridge distance in the -Mg^{2+} system in both the NPT_1 and NPT_2 simulations (Fig. 2.24 and 2.25). This suggests that Arg 346$_a$::Glu 347$_b$ is susceptible to removal of Mg^{2+}, even though the interaction remained intact.

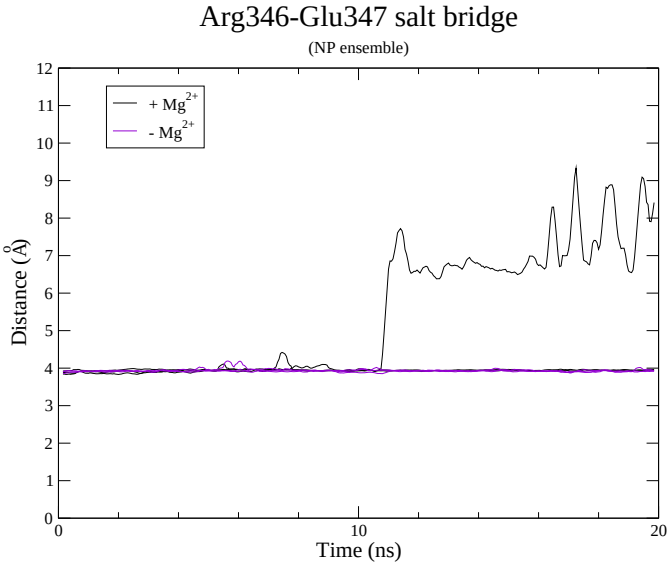


Figure 2.23: Effect of removing Mg^{2+} on the Arg 346$_a$::Glu 347$_b$ salt-bridge in the NP ensemble. One interaction is observed to break during the simulation with Mg^{2+} included. The plots were averaged using a sliding window of 500 frames (250 fs).

As described above, visual inspection of the homology models suggested a further interaction at Glu 295$_a$::Arg 404$_b$. While not fully formed in the homology models, the salt-bridge distance did adopt standard values (±4 Å) during minimisation and heating of the systems. The integrity of this interaction was found to be more susceptible to removal of Mg^{2+} than Arg 346$_a$::Glu 347$_b$. In the -Mg^{2+} systems, the salt-bridge was broken in half the possible cases over both the NP and NPT ensembles. In the NP ensemble the interaction is broken between two of the monomer pairs (between chain A and B, and chain B and C) by the end of 20 ns. The third interaction (chain C and A) was transiently broken (Fig. 2.26). In the NPT_1 ensemble two salt bridges (chain A and B, chain B and C) were broken by the end of 50 ns of simulation (Fig. 2.27), while one interaction (chain A and B) was broken by the end of 50 ns in the NPT_2 ensemble (Fig. 2.28). Since the side-chains would be freer to move in the broken interactions instead of being locked in by the salt-bridge interaction, this causes

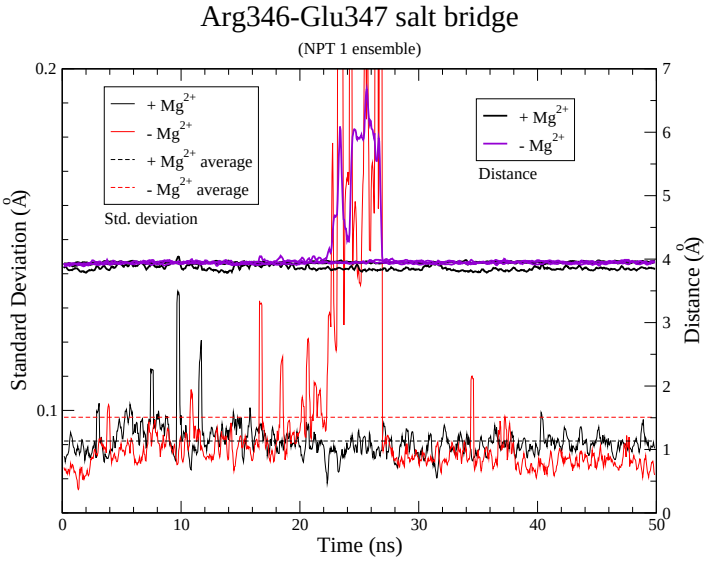


Figure 2.24: Effect of removing Mg^{2+} on the Arg 346_a::Glu 347_b salt-bridge in the NPT_1 ensemble. One interaction is observed to break briefly during the simulation without Mg^{2+}. The distances for all three salt bridge distances are overlayed: thick black (with Mg^{2+}) and thick purple (without Mg^{2+}) lines. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

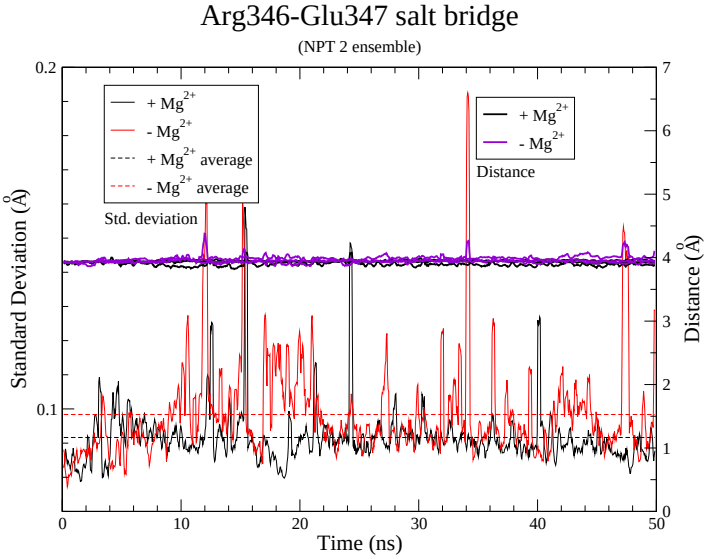


Figure 2.25: Effect of removing Mg^{2+} on the Arg 346_a::Glu 347_b salt-bridge in the NPT_2 ensemble. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

a marked increase in the standard deviation of the Glu 295_a::Arg 404_b distances in the -Mg^{2+} systems (results not shown).

In the NP ensemble, the temperature of the system was initially 310 K, but was not kept constant. During the sampling run the temperature gradually increased to approximately 332 K (Fig. 2.29). It is usual to apply some means to keep the temperature constant (isothermal ensemble), however, due to a mistake in the input files, this was omitted. Nonetheless, the results from the NP ensemble are still interesting and somewhat serendipitous. By sampling at higher temperatures, it is possible to overcome barriers in the energy landscape faster than at lower temperatures. The increase in temperature apparently accelerates the effects of removing Mg^{2+}. In the NPT simulations, only one Glu 295_a::Arg 404_b interaction is broken after 20 ns, whereas in the NP ensemble with increasing temperature all three have been broken before 20 ns. The effect of the increasing temperature is also reflected in the RMSD, which is more pronounced and more rapid in the NP ensemble. The increasing temperature may be detrimental, however, as reflected by the breaking one of the Arg 346_a::Glu 347_b

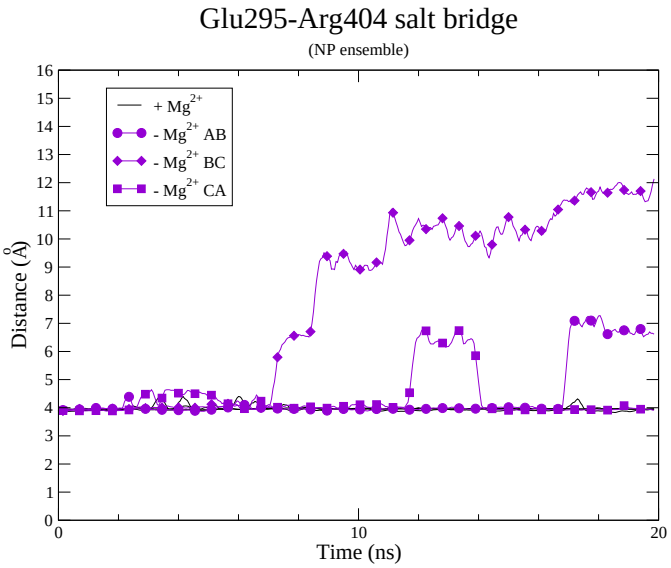


Figure 2.26: Effect of removing Mg^{2+} on the Arg 295_a::Glu 404_b salt-bridge in the NP ensemble. Without Mg^{2+} two interactions are broken at the end of the simulation, while the other is transiently broken. The three arginase monomers are arbitrarily designated A, B and C, with ◯, ◊ and □ used to indicate the three salt bridges. The plots were averaged using a sliding window of 500 frames (250 fs).

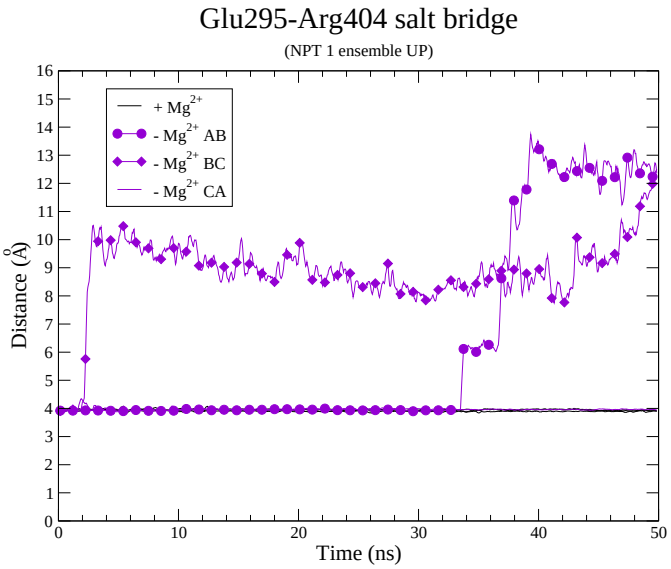


Figure 2.27: Effect of removing Mg^{2+} on the Arg 295_b::Glu 404_b salt-bridge in the NPT_1 ensemble. Without Mg^{2+} two interactions are broken at the end of the simulation. The three arginase monomers are arbitrarily designated A, B and C, with ◯ and ◊ used to indicate the two broken salt bridges. The plots were averaged using a sliding window of 500 frames (250 fs).

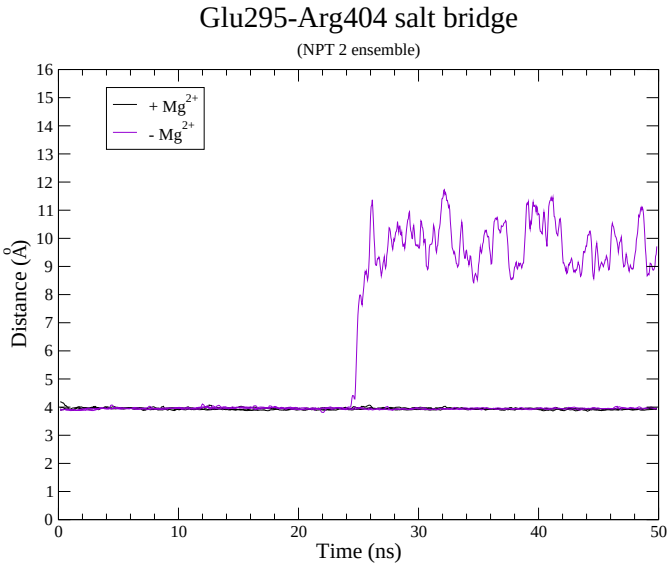


Figure 2.28: Effect of removing Mg^{2+} on the Arg 295_a::Glu 404_b salt-bridge in the NPT_2 ensemble. Without Mg^{2+} two interactions are broken at the end of the simulation. The three arginase monomers are arbitrarily designated A, B and C. The plots were averaged using a sliding window of 500 frames (250 fs).

interactions in the NP ensemble with Mg^{2+}. For this reason, all subsequent simulations were carried out in the NPT ensemble.

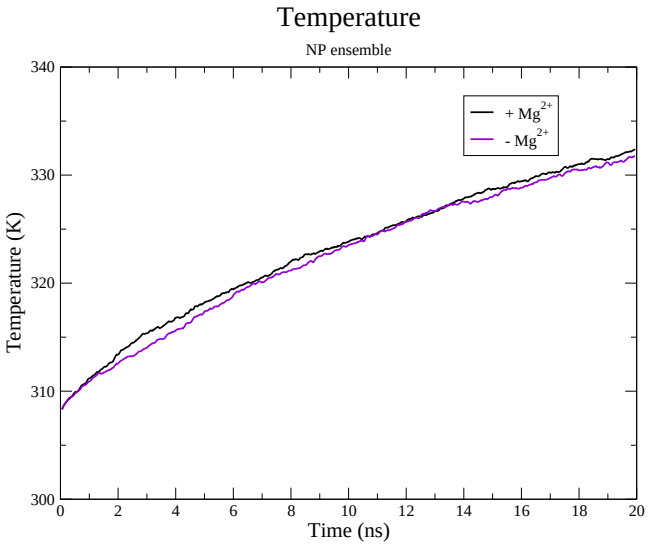


Figure 2.29: Temperature increase in the NP ensemble. When no pressure control was implemented the temperature of the system increased by the same amount in both systems with and without Mg^{2+}. The plots were averaged using a sliding window of 500 frames (250 fs).

Müller *et al.* (2005) demonstrated that mutants that abolish metal binding also disturb the trimer and that mutating the conserved interaction between Arg 346$_a$::Glu 347$_b$ results in inactive monomers. As a consequence of the simulations with and without Mg^{2+} it was decided to test the possibility of Glu 295$_a$::Arg 404$_b$ interacting experimentally. The experimental work was performed by I. B. Müller of the Bernhard-Nocht institute of Tropical Medicine, Hamburg, Germany. It was reasoned, that if Glu 295$_a$::Arg 404$_b$ is involved in this dependency, enzyme activity and trimer formation should also be susceptible to mutation. *P. falciparum* arginase was found to be more susceptible to mutations introduced at Glu 295 than Arg 404 (Table 2.4). Mutating Glu 295 to Ala or Arg considerably reduces enzyme activity by 96% and 73% respectively. Mutating Arg 404 to Ala has a less dramatic result (54% less active) but is still substantial. The double mutation of Glu 295 Ala/Arg 404 Ala also essentially inactivates the enzyme (95% less active). However, the single mutations of Glu 295 Arg and Arg 404 Ala leads to altered K_m values of 146 mM and 45 mM for L-arginine, respectively, which is up to 11-fold higher compared to the wild type arginase. The elution profile of all mutants analysed by gel filtration revealed monomeric forms, except for Glu 295 Ala, which is partially trimeric (Figure 2.30). In contrast, trimer formation is more susceptible to mutation of Arg 404, compared to enzyme activity.

Table 2.4: Effect of Glu 295$_a$::Arg 404$_b$ salt-bridge mutations (E295A, E295R, R404A and E295A/R404A) on arginase activity; standard error of the mean (SEM). ND: not detectable.

Mutant	V_{max} ($\mu mol.min^{-1}.mg^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m, ($mM^{-1}.s^{-1}$)
Wild type	31	13	24.8 (100%)	1.9 (100%)
Glu 295 Ala	1.3 ± 0.3	ND	1.0 ± 0.2 (4%)	ND
Glu 295 Arg	8.4 ± 0.9	146 ± 6	6.7 ± 0.7 (27%)	0.03 (1.6%)
Arg 404 Ala	14.3 ± 0.9	45 ± 3	11.4 ± 0.7 (46%)	0.25 (13%)
Glu 295 Ala/ Arg 404 Ala	1.6 ± 0.1	ND	1.3 ± 0.1 (5%)	ND

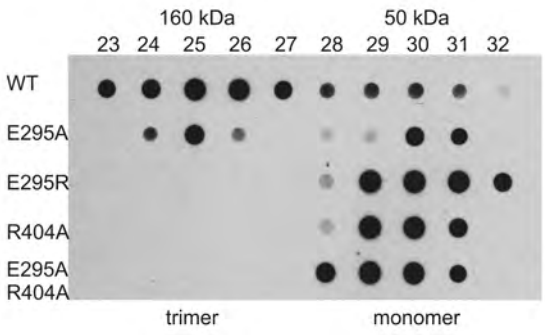


Figure 2.30: Effect of Glu 295$_a$::Arg 404$_b$ salt-bridge mutations (E295A, E295R, R404A and E295A/R404A) on trimer formation.

Thus, the existence of the inter-monomer salt-bridge at Glu 295$_a$::Arg 404$_b$ was confirmed by site-directed mutagenesis of the recombinant enzyme. All mutants tested promoted trimer dissociation, with incomplete dissociation for Glu 295 Ala. This contrasts with Glu 295 Arg which gave complete dissociation. Mutating Glu 295 to Arg is expected to be more drastic compared to Ala, because this would potentially introduce a positive charge in the vicinity of the Glu 295$_a$::Arg 404$_b$ interaction, and thus be more likely to disturb the trimer. Interestingly, this mutation leads to active but less efficient monomers with an increased K_m value of 146 mM for L-arginine, indicating altered substrate binding. In contrast, mutating Glu 295 to Ala reduced the activity to 4% of the wild-type enzyme but with its trimeric conformation partially retained. The K_m value for the Glu 295 Ala mutant was not measurable because it was not saturated up to 200 mM arginine. Mutating Arg 404 to Ala abolished trimer formation. However, this mutant enzyme shows 46% activity (as *kcat*) and 13% efficiency (as $kcat.K_m^{-1}$) and its K_m value is approximately three-fold increased compared to the wild-type enzyme. This result is similar to the previously reported behaviour of the rat liver arginase I Arg 308 mutants, which, as monomers, still had a residual activity of 41% and an efficiency in the range 13–17% (Lavulo *et al.*, 2001). Size-exclusion chromatography

therefore suggests that certain mutations abolish trimerisation but the enzymatic data suggests that trimerisation is not absolutely necessary for activity. However, the possibility that a weakened trimer can form under enzyme assay conditions cannot be excluded. Such a possibility is suggested by rat arginase, where the Arg 308 Lys mutant is apparently active as a monomer, but nonetheless crystallises as a trimer (Lavulo *et al.*, 2001). Although it has been demonstrated that disturbing the oligomer via the conserved Arg 346_a::Glu 347_b interaction largely inactivates the enzyme, it still has 10% residual activity (Müller *et al.*, 2005). The results of the Arg 404 Ala mutation indicates that it is possible to produce active monomers and furthermore, that certain mutations can partially compensate for induced structural instability of monomerisation by long range allosteric effects. Although there is a dependency between trimer formation and enzyme activity, these results indicate that it is not complete. This incompleteness was suggested by previous results where mutating His 193 in the active site also results in an inactive trimer (Müller *et al.*, 2005) as was also found for the Glu 295 Ala mutation in the present study. Nonetheless, formation of the predicted salt-bridge is apparently necessary for trimer formation.

The effect of mutating the Glu 295_a::Arg 404_b interaction was followed up by molecular dynamics of the mutants. During simulation, the Arg 346_a::Glu 347_b salt-bridge remains largely intact with the introduction of mutations into the Glu 295_a::Arg 404_b salt-bridge. The exceptions to this are the temporary breaking of one Arg 346_a::Glu 347_b interaction for about 10 ns in the Arg 404 Ala mutant (Fig. 2.34) and for about 5 ns in the Glu 295 Arg mutant (Fig. 2.33). In all other cases the salt-bridge distance remains at 4 Å. However, the fluctuation around the 4 Å is affected by introduction of alanine at Glu 295. In both the Glu 295 Ala and Glu 295 Ala/Arg 404 Ala mutants the average standard deviation decreases compared to the wild type (Fig. 2.31, 2.32). In contrast the introduction of Glu 295 Arg results in an increased standard deviation (Fig. 2.33), as does the Arg 404 Ala mutant. In both cases this is largely the result of the temporary breaking of one interaction (Fig. 2.34, Fig. 2.33). The introduction of the Glu 347 Gln mutant has a profound effect on Glu 295_a::Arg 404_b. All three of the interactions are broken, with one of these already disrupted during the heating stage (Fig. 2.34).

The Arg 346_a::Glu 347_b interaction can potentially interact with Glu 295 (Fig. 2.9). This is confirmed by the mutation of Glu 295 to neutral Ala which is less likely to interact strongly with either of the charged residues of Arg 346_a::Glu 347_b. In both simulations with this mutation (Glu 295 Ala and Glu 295 Ala/Arg 404 Ala) there is a lower standard deviation for the Arg 346_a::Glu 347_b distance compared to the wild type. In contrast, Glu 295 Arg causes in increase in the standard deviation compared to the wild type. Like Glu 295, Arg 295 would also be expected to interact more strongly with Arg 346_a::Glu 347_b. It is suggested that the longer sidechain of Arg compared to Glu could potentially bring it in closer contact with Arg 346_a::Glu 347_b, thus resulting in the increased effect on that interaction. Interestingly, the mutation of Arg 404 Ala also increases the standard deviation of the Arg 346_a::Glu 347_b distance. This residue is to far to make physical contact. While this might be a chance

effect, this result is in agreement with experiment which suggests that trimer formation is more susceptible to mutation at Arg 404 than Glu 295 (Fig. 2.30). The disruption of Glu 295_a::Arg 404_b by the Glu 347 Gln mutation further strengthens the suggestion that Glu 295_a::Arg 404_b is a required inter-monomer interaction, and that disruption of the trimer by mutations at Glu 347 or metal removal are mediated by this salt-bridge.

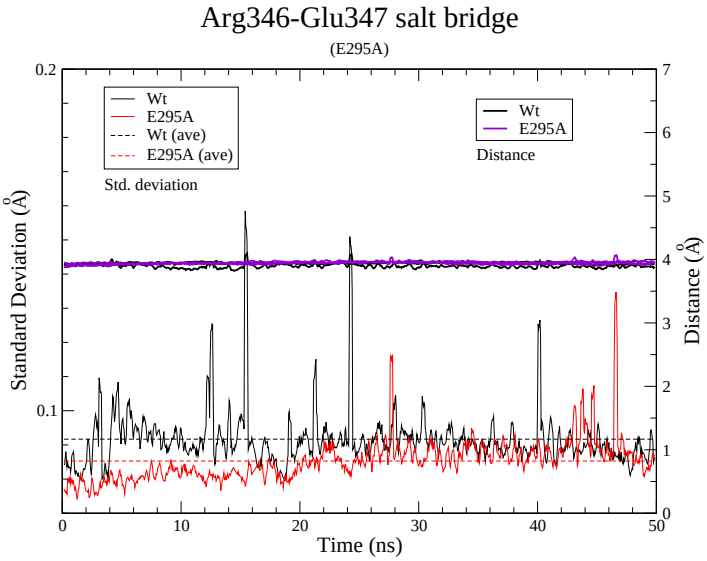


Figure 2.31: Effect of the Glu 295 Ala mutation on the Arg 346_a::Glu 347_b salt-bridge. A higher standard deviation is observed for the wild type. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

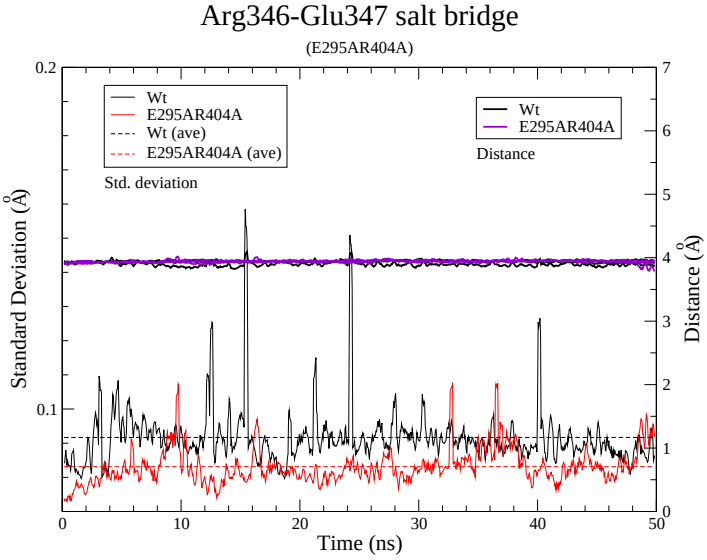


Figure 2.32: Effect of the Glu 295 Ala/Arg 404 Ala mutation on the Arg 346_a::Glu 347_b salt-bridge. A higher standard deviation is observed for the wild type. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

2.4.3.6 Co-ordination geometry of Mg^{2+}

It is expected that disturbing the interactions involved in trimer formation mediate their effects via the co-ordination of Mn^{2+} in the active site, which is required for the arginase chemistry. This is reflected by the increased equilibrium RMSD during MD, which should ultimately translate into lost co-ordination of Mn^{2+} in the active site. The loss of Mn^{2+} under such conditions, however, has yet to be observed directly. Therefore, the effect of the mutations on Mg^{2+}co-ordination during MD was analysed.

In the arginase active site, highly conserved residues are involved in a specific co-ordination pattern for the binuclear Mn^{2+} cluster. In rat I arginase the more deeply buried ion (Mn_A^{2+}) is

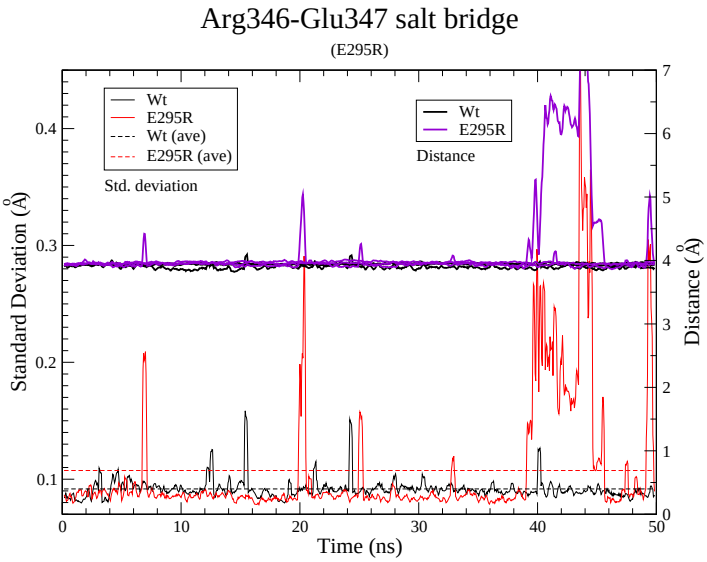


Figure 2.33: Effect of the Glu 295 Arg mutation on the Arg 346_a::Glu 347_b salt-bridge. A higher standard deviation is observed for the mutant. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

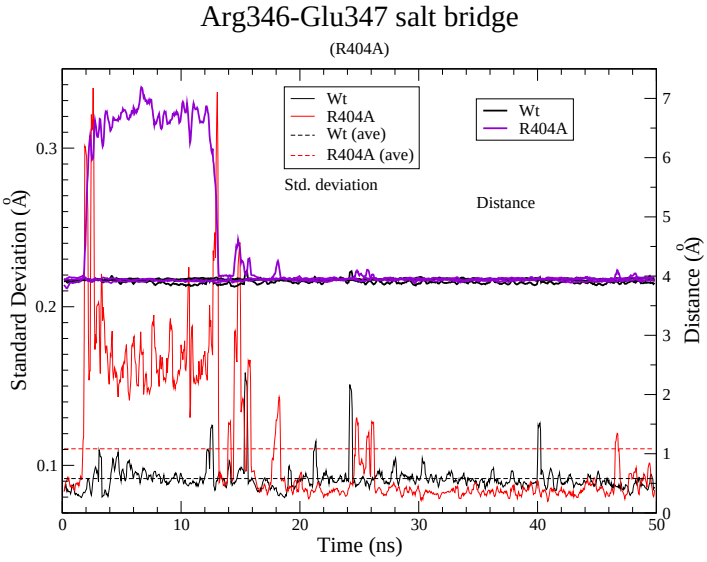


Figure 2.34: Effect of the Arg 404 Ala mutation on the Arg 346_a::Glu 347_b salt-bridge. A higher standard deviation is observed for the mutant. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

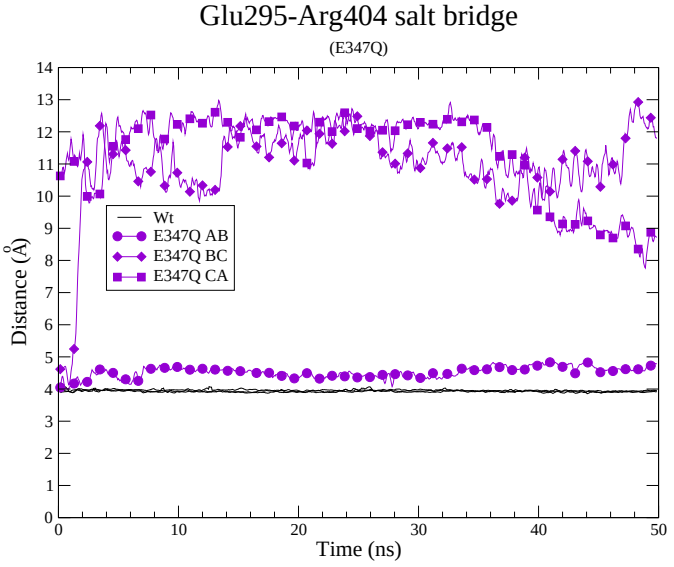


Figure 2.35: Effect of the Glu 347 Gln mutation on the Glu 295_a::Arg 404_b salt-bridge. The mutations disrupts Glu 295_a::Arg 404_b. The distance plots were averaged using a sliding window of 500 frames (250 fs). The standard deviation calculated with a sliding window of 500 and averaged for the simulations with Mg^{2+} (thin black) without Mg^{2+} (thin red). The average standard devation over the entire simulation is indicated with dashed lines.

co-ordinated by His 101, Asp 124, Asp 128, Asp 232 and the bridging solvent in a square pyramidal geometry. The respective residues in *P. falciparum* are His 193, Asp 216, Asp 220 and Asp 323. The second metal, Mn_B^{2+} is co-ordinated by His 126, Asp 124, Asp 232, Asp 234 and the bridging solvent in a distorted octahedral geometry (His 218, Asp 216, Asp 323, Asp 325 in *P. falciparum*).

During the simulations, the conformations adopted by the co-ordinating residues did not entirely conform to known crystal structures from homologues. The most notable difference is Asp 323, which is expected to form a monodentate bridging interaction between the two ions. During the simulations, it formed a bidentate bridge instead (Fig. 2.36). All other expected co-ordinating atoms were oriented close enough to interact with the ions. The only other missing interaction was that of the bridging OH^- as no attempt was made to introduce the bridging solvent molecule. The Mg^{2+}-Mg^{2+} distance was also about 0.6 Å greater than the known Mn^{2+}-Mn^{2+} distance. The larger distance is partly due to the inability of the software to recognise co-ordination chemistry natively as well as the larger van der Waals radius of Mg^{2+} compared to Mn^{2+}. The Mg^{2+} ions nonetheless remained in the active site during the simulations and restricted the movement of the interacting ligands.

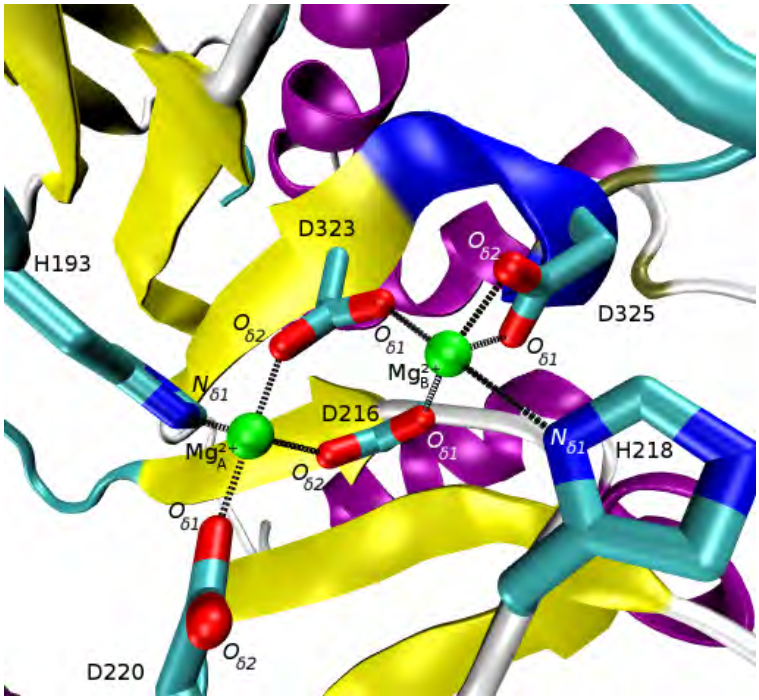


Figure 2.36: General co-ordination pattern of Mg^{2+} in the NPT_2 ensemble in chain A after 40 ns. The bridging OH^- is absent. Mg_A^{2+} is co-ordinated in an incomplete trigonal bipyramidal geomentry, while Mg_B^{2+} is co-ordinated in an incomplete octahedral geometry. The missing ligand is an OH^- which could not be included for modeling due to force-field restrictions. The exact geometry also differ slightly compared to experimental structures. The presence of Mg^{2+} still restricts ligand movement, however.

The stability of the active site was observed by monitoring the inter-atom distances between the interacting residues and Mg^{2+}. Four interactions occur between Mg_A^{2+} and His $193N_{\delta 1}$, Asp $220O_{\delta 1/\delta 2}$, Asp $216O_{\delta 1/\delta 2}$ and Asp $323O_{\delta 1/\delta 2}$. Five interactions occur between Mg_B^{2+} and His $218N_{\delta 1}$, Asp $216O_{\delta 1/\delta 2}$ Asp $323O_{\delta 1/\delta 2}$, Asp $325O_{\delta 1}$ and Asp $325O_{\delta 2}$. For the acidic residues Asp 323, Asp 216 and Asp 220 generally only one of the chemically equivalent carboxyl oxygens ($O_{\delta 1}$ or $O_{\delta 2}$) interacts with an Mg^{2+} ion. Although these oxygens may potentially exchange places in interacting with Mg^{2+}, this was never observed during any of the solvated simulations. The interacting carboxyl oxygens generally maintained a constant distance of ± 1.8 Å with the Mg^{2+} ion, while the non-interacting oxygens displayed

greater variability. During the NPT_2 simulation an apparent partial exchange of position was observed for Asp 220 with Mg_A^{2+} in chain C where both $O_{\delta 1}$ and $O_{\delta 2}$ distances reach ± 1.9 Å. The active sites of the NPT simulations were generally stable during 50 ns simulation, apart from the altered conformation of Asp 220 described above, as well as a slight increase for the His 193 $N_{\delta 1}$-Mg_A^{2+}interaction, also in chain C. In the NP simulation, there was more instability. The distances between Mg_A^{2+} and both $O_{\delta 1}$ and $O_{\delta 2}$ of Asp 216 increased to about ± 5.8 Å from ± 1.8 Å by the end of 20 ns in chain C (Fig. 2.37). Furthermore, the distance between His 218 $N_{\delta 1}$ and Mg_B^{2+} increased to ± 7 Å (Fig. 2.38). The Asp 216 and His 218 interactions with their respective Mg^{2+} were thus essentially broken. The greater instability during the NP simulation is thought to be due to the increase in temperature that occurred (explained below). It is expected that the interactions with Mg^{2+} in the active site should remain stable, since the bi-nuclear cluster is required for enzyme activity.

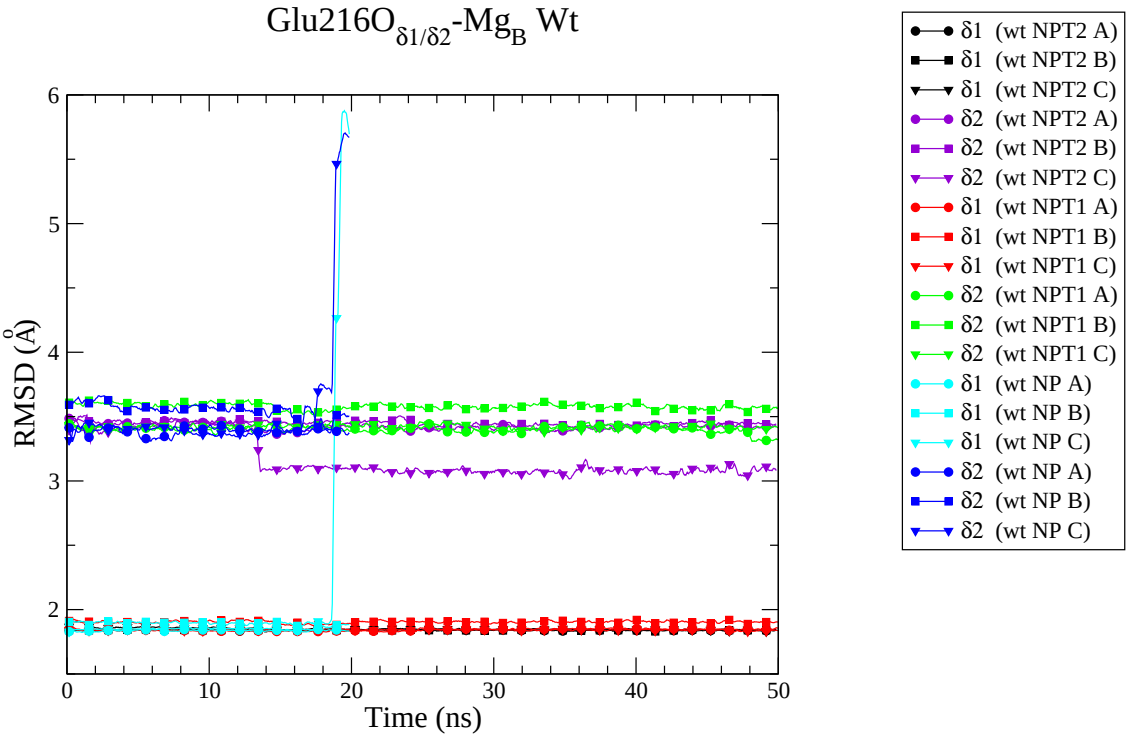


Figure 2.37: Distance between Asp 216 $O_{\delta 1/\delta 2}$ and Mg_B^{2+} in wild type arginase in the NP, NPT_1 and NPT_2 simulations. Both carboxyl oxygens ($\delta 1/\delta 2$) are included. Typically only one carboxyl O interacts with Mg^{2+} at a distance of ± 1.8 Å, leaving the other carboxyl O to vary at ± 3.4 Å. Most interactions remained stable. One residue interaction was broken in chain C of the NP simulation. Pairs of carboxyl O are indicated for chains A (○), B (□) and C (▽).

When simulating without Mg^{2+} removed, considerable movement is observed in the co-ordinating residues in all three simulations (NP, NPT_1 and NPT_2). However, in the NP simulation there is more more movement of the co-ordinating residues with Mg^{2+} compared to the NPT simulations. This is likely to be due to the increased temperature during the simulation (Fig. 2.39).

Because the presence of Mg^{2+} was able to stabilise the co-ordinating residues by electro-

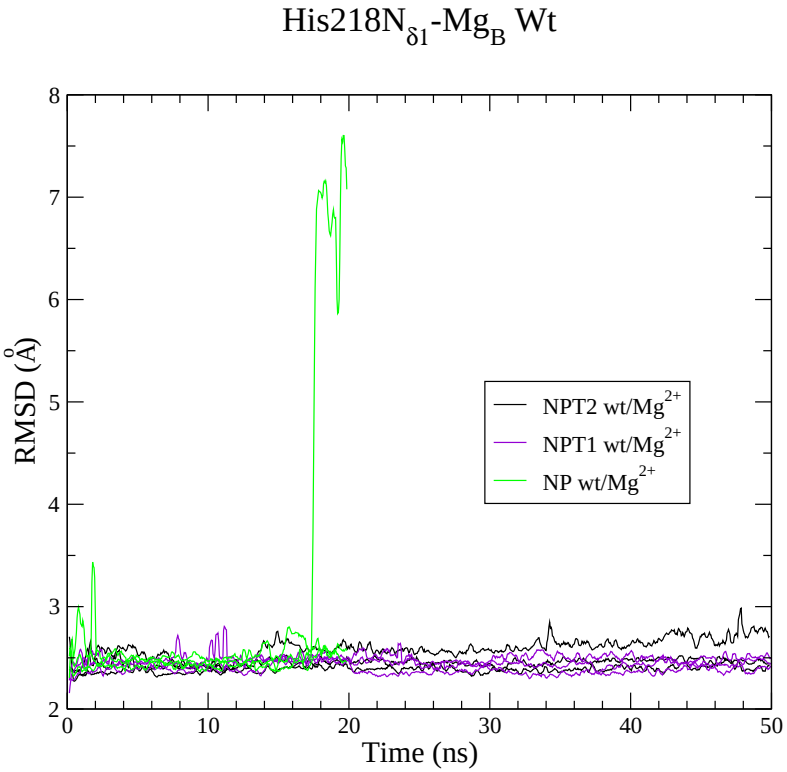
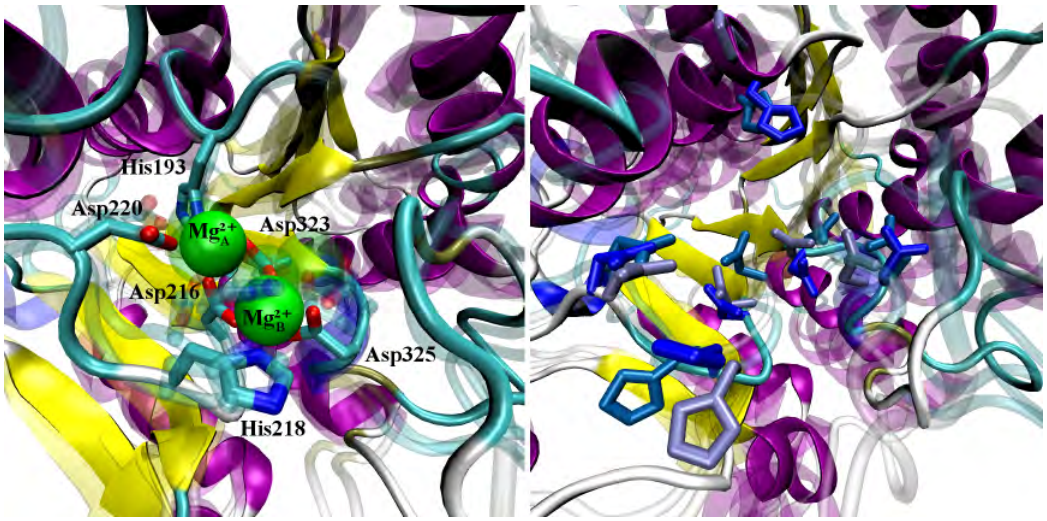


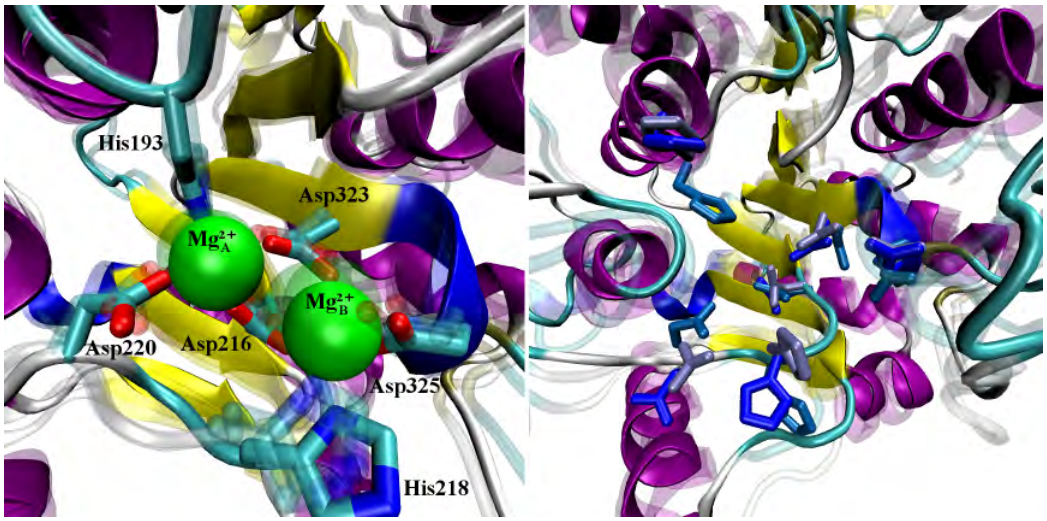
Figure 2.38: Distance between His 218$N_{\delta1}$ and Mg_B^{2+} in wild type arginase in the NP, NPT_1 and NPT_2 simulations. Most interactions remained stable. One interaction is broken in chain C of the NP simulation. The plots for chains A, B and C are overlaid.

static interactions alone, this approach appears viable for investigating the structural metal dependency. This suggests that structural metal dependency involves free movement of the metal co-ordinating residues. The extra stability of the active sites in the NPT ensembles compared to the NP ensemble, suggests that this is the more appropriate type of simulation.

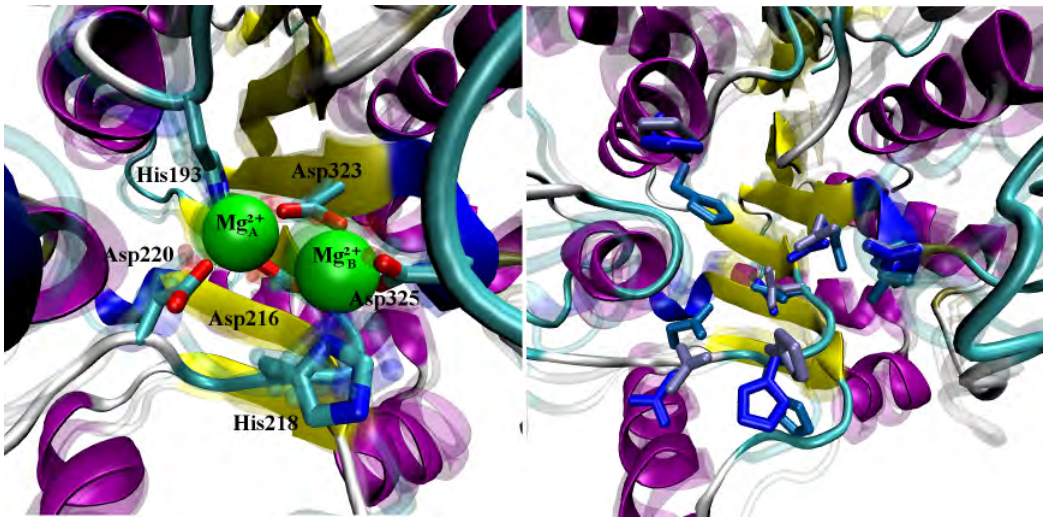
During the simulation of the various mutants, disturbances were observed in the active site compared to the wild type simulations. In three out of five mutants there were large deviations in the interaction distances compared to the standard distances observed during the wild type simulations. No interactions were lost during the simulation of Arg 404 Ala and Glu 295 Ala/Arg 404 Ala. In the Glu 295 Ala simulation the distance between Asp 216$O_{\delta1/\delta2}$ and Mg_B^{2+} increases to 5.0-5.5 Å (from ± 1.8 Å) for about 6 ns (Fig. 2.40) and the distance between His 218$N_{\delta1}$ and Mg_B^{2+}increases to 4.5-7.5 Å (from ± 1.8 Å) for about 8 ns (Fig. 2.41). In Glu 295 Arg and Glu 347 Gln mutants more drastic increases in the interaction distances were observed. In Glu 295 Arg the distance between Asp 216$O_{\delta1/\delta2}$ and Mg_B^{2+} increases to ± 4.5-6 Å for about 30 ns in chain B (Fig. 2.42). Furthermore – also in chain B – the distance between His 218$N_{\delta1}$ and Mg_B^{2+} increases from ± 5 Å to ± 8 Å during 40 ns (Fig. 2.43). The distance had already increased to ± 5 Å during the heating stage. The greatest number of disturbances are observed for Glu 347 Gln. A number of these are observed early in the production run, suggesting that mutating Glu 347 to Gln introduces a significant disturbance. Firstly, the distance between Asp 323$O_{\delta1/\delta2}$ and Mg_A^{2+} increases to about ± 3.8



(a) NP



(b) NPT_1



(c) NPT_2

Figure 2.39: Destabilisation of active site metal co-ordinate residues. All three chains within each of the $+Mg^{2+}$(left) and $-Mg^{2+}$(right) from the NP, NPT_1 and NPT_2 simulations were superimposed. Two out of the three chains are represented as transparent, except for the co-ordinating residues for the $-Mg^{2+}$simulations, which are indicated in different shades of blue. When Mg^{2+} is removed there is greater movement of the co-ordinating residues compared to when $Mg^{2+}$is present.

Å and ± 5.1 Å for each of the sidechain carboxyl oxygens for 40 ns in chain C (Fig. 2.44). Secondly, the interaction between His 218 $N_{\delta 1}$ and Mg_B^{2+} increases to ± 6-8 Å for chains B and C for about 39 ns each (Fig. 2.45). The interactions for which there were no significant deviations are included in Appendix A.

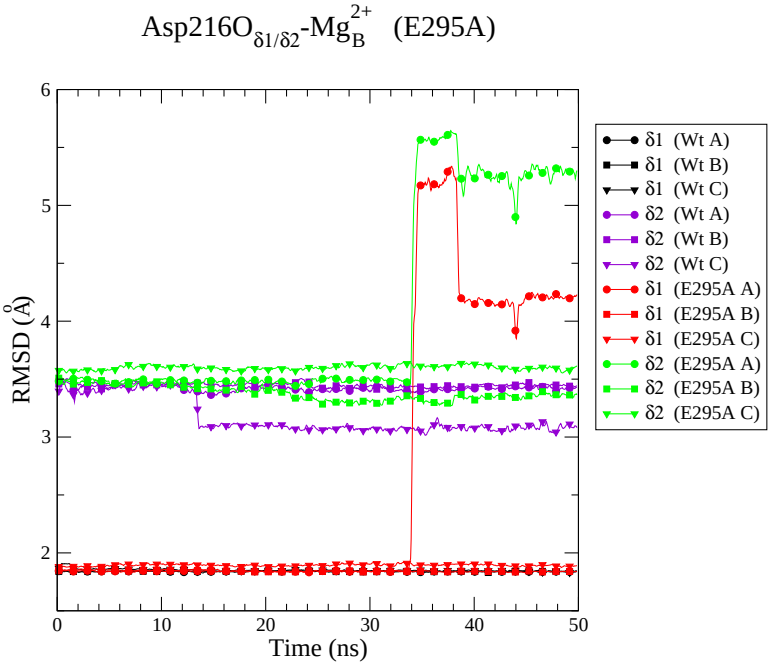


Figure 2.40: Interaction between Asp 216 O and Mg_B^{2+} in *pf* Arg Glu 295 Ala compared to wild type. Both carboxyl oxygens ($\delta 1/\delta 2$) are included. Pairs of carboxyl O are indicated for chains A (○), B (□) and C (▽). In the Glu 295 Ala mutant one interaction is broken (chain C) for the last six ns.

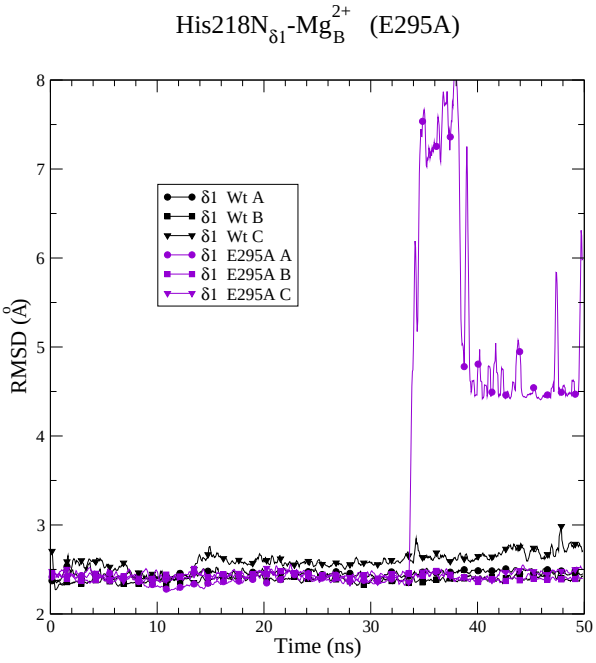


Figure 2.41: Interaction between His 218 $N_{\delta 1}$ and Mg_B^{2+} in *pf* Arg Glu 295 Arg compared to wild type. Plots are indicated for chains A (○), B (□) and C (▽). In the Glu 295 Ala mutant one interaction is broken (chain C) for the last eight ns.

It has been demonstrated certain mutations to *Pf* Arg located away from the active site and partaking in trimer formation, compromise activity. Furthermore, mutations that abolish metal binding, or removal of the catalytic metal, compromise activity (Müller *et al.*, 2005). Therefore, it is suggested that mutations to the inter-monomer salt-bridges mediate their effects by causing the loss of metal binding. The simulations with of the *Pf* Arg mutants suggests this may be the case. During the simulation of certain mutants, there was partial

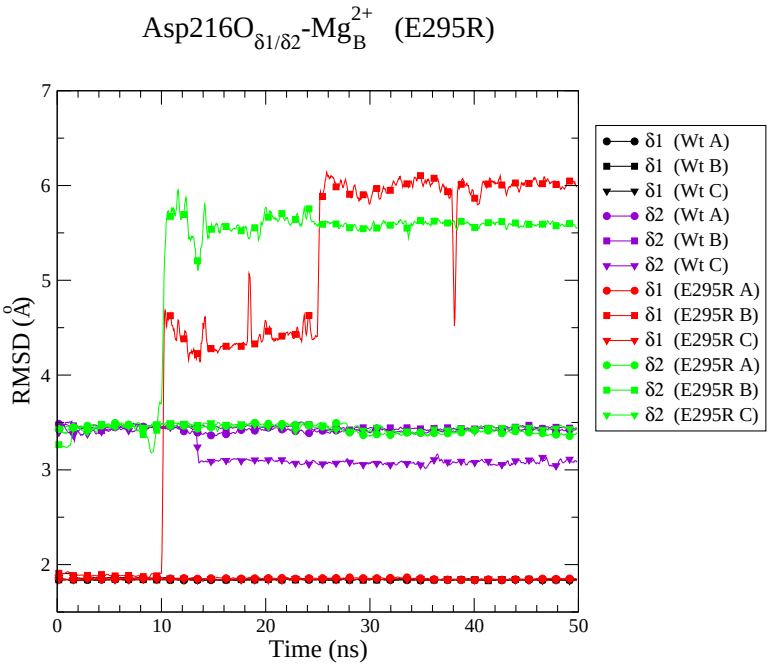


Figure 2.42: Interaction between Asp 216 O and Mg^{2+}_{B} in *pf*Arg Glu 295 Arg compared to wild type. Both carboxyl oxygens ($\delta1/\delta2$) are included. Pairs of carboxyl O are indicated for chains A ($\circ$), B ($\square$) and C ($\triangledown$). In the Glu295Arg mutant one interaction is broken (chain B) for the last 30 ns.

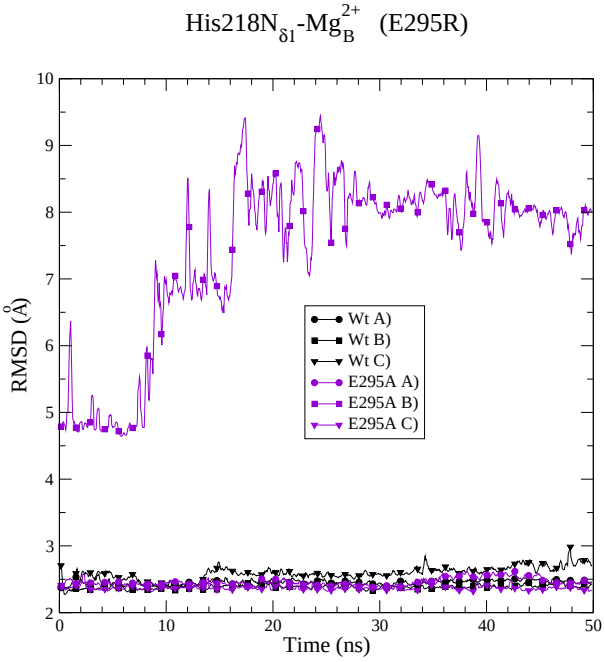


Figure 2.43: Interaction between His 218 $N_{\delta1}$ and Mg^{2+}_{B} in *pf*Arg Glu 295 Ala compared to wild type. Both carboxyl oxygens ($\delta1/\delta2$) are included. Pairs of carboxyl O are indicated for chains A ($\circ$), B ($\square$) and C ($\triangledown$). In the Glu 295 Arg mutant one interaction is broken (chain B) for 40 ns.

loss of the interactions expected to maintain the metals in the active site. Loss of these interactions is likely to lead to loss of metal binding and thus activity. During the simulations, the greatest disturbance to interactions with Mg^{2+} was observed for Glu 347 Gln, Glu 295 Arg and Glu 295 Ala, whereas little disturbance was observed for mutations to Arg 404 (Arg 404 Ala and Glu 295 Ala/Arg 404 Ala). This is in partial agreement with experimental results concerning these mutations. Loss of activity is observed for all these mutations, with the least reported loss for Arg 404 Ala (Table 2.4) and the greatest loss for mutations to Glu 295 or Glu 347 (Müller *et al.*, 2005). However, although the double mutant Glu 295 Ala/Arg 404 Ala results in 95% loss of activity, simulations of this mutant did not reveal any disturbance to the active site. Furthermore, although there were disturbances to the active site during the Glu 295 Ala simulation, these occurred only within one chain for a relatively brief period.

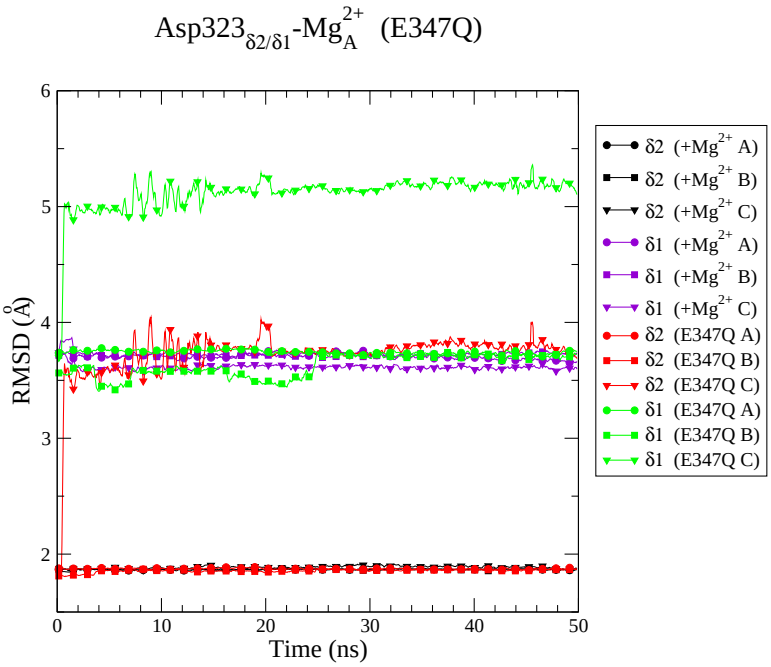


Figure 2.44: Interaction between Asp 323 O and Mg_A^{2+} in *pf*Arg Glu 347 Gln compared to wild type. Both carboxyl oxygens ($\delta 1/\delta 2$) are included. Pairs of carboxyl O are indicated for chains A ($\circ$), B ($\square$) and C ($\triangledown$). In the Glu347Gln mutant one interaction is broken (chain C) for the last 30 ns.

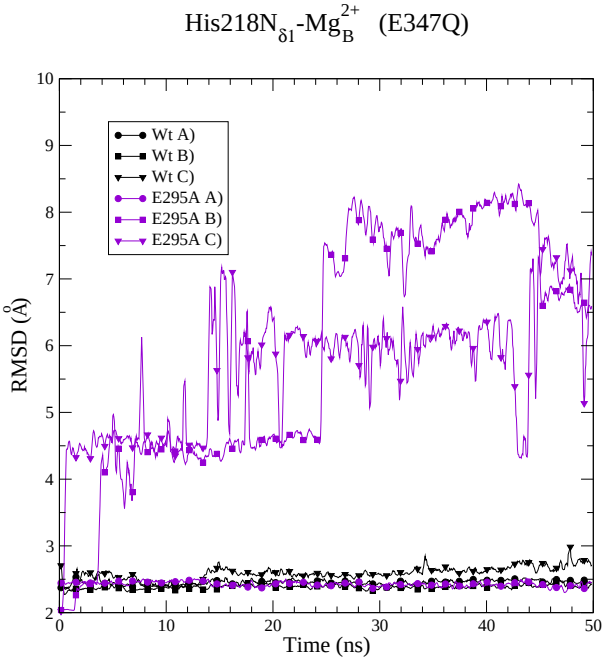


Figure 2.45: Interaction between His 218 N and Mg_B^{2+} in *pf*Arg Glu 347 Gln compared to wild type. Plots are indicated for chains A ($\circ$), B ($\square$) and C ($\triangledown$). In the Glu347Gln mutant two interactions are broken (chains B, C) for about 39 ns each.

Thus, the simulations performed do not completely capture what is observed experimentally. The times simulated were relatively short (50 ns) and thus may need to be extended to obtain agreement with experiment.

2.4.3.7 Movement per residue

In order to determine whether any part of the protein reproducibly moved more upon removal of the active site metals, the RMSD per residue was analysed. This was carried out in VMD by aligning the whole protein of each frame of the sampling run to the first frame. The change in RMSD over each residue was determined for each frame and summed. This value was then averaged by the number of frames. In the NP simulation, certain portions of the protein that form part of the inter-monomer interfaces were observed to have a higher RMSD

per residue (Fig. 2.46). However, this was not reproducible in the NPT simulations.

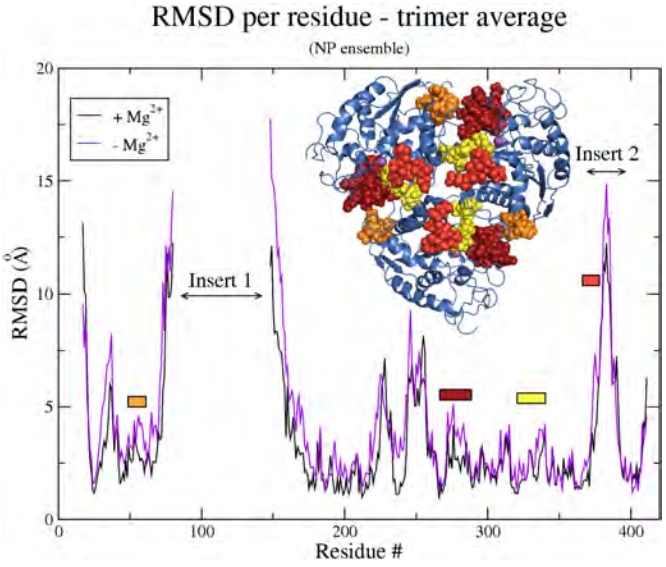


Figure 2.46: RMSD per residue for the NP simulation. When Mg^{2+} is removed certain regions are observed to move more than with Mg^{2+}. Among these regions are inter-monomer contacts indicated in the structure. Each chain was superimposed (all protein atoms) independently. The average for all three chains in each system (with and without Mg^{2+}) is shown. All frames from the simulation were used for the calculation.

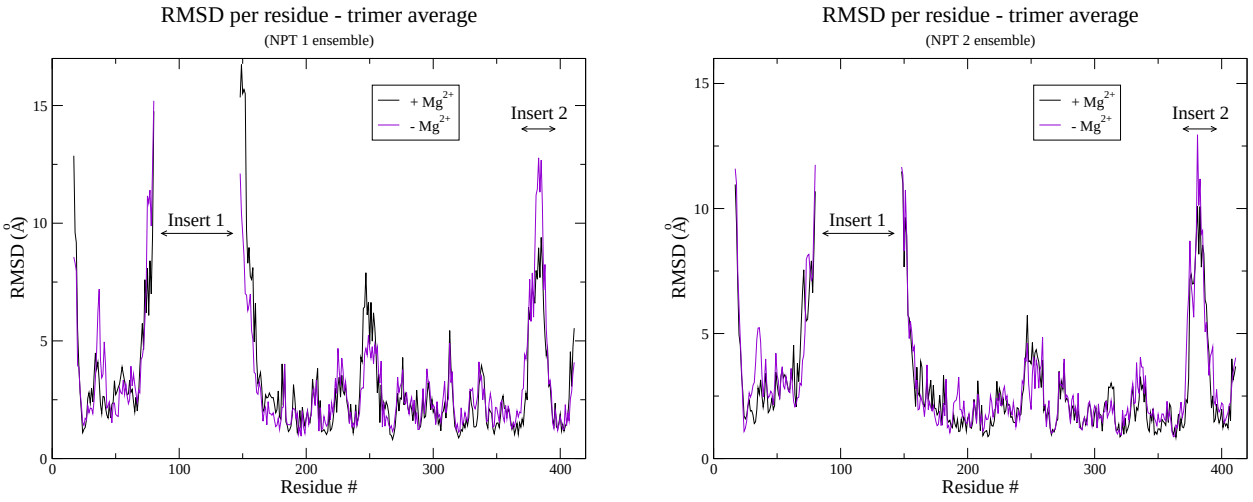


Figure 2.47: RMSD per residue for the NPT simulations. Compared to the NP simulation the same increase in movement of certain interacting regions without Mg^{2+} is not observed. Each chain was superimposed (all protein atoms) independently. The average for all three chains in each system (with and without Mg^{2+}) is shown. All frames from the simulation were used for the calculation.

The RMSD per residue results are generally mirrored when calculating the root mean square fluctuation (RMSF) instead, which is the standard deviation of the RMSD. In the NP simulation the RMSF per residue results mirror that of RMSD per residue (Fig. 2.48). RMSF of the $C_\alpha$atoms gives less noisy plots compared to RMSD per residue, and it becomes possible to observe increased movement in certain regions in the NPT simulations (Fig. 2.49). The increased movement of certain interface contacting regions observed in the NP simulation is partially repeated in the NPT simulations, however. There is increased movement for insert 2 in all three simulation, in agreement with increased movement of insert 2 towards the interface regions without Mg^{2+}.

The increase in RMSD observed in the NP ensemble but not repeated for NPT could be due to the increase in temperature that occurred in the NP simulation. Furthermore,

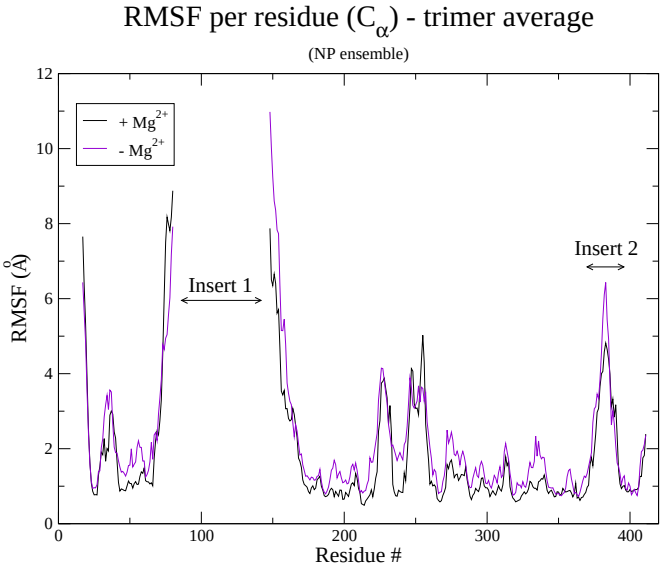


Figure 2.48: Root mean square flucation (RMSF) or standard deviation of RMSD per residue for the NP simulation (C_α atoms). This generally mirrors the RMSD per residue results. Each chain was superimposed (all protein atoms) independently. The average for all three chains in each system (with and without Mg^{2+}) is shown. Every 100th frame was used (400 frames total) for the calculation.

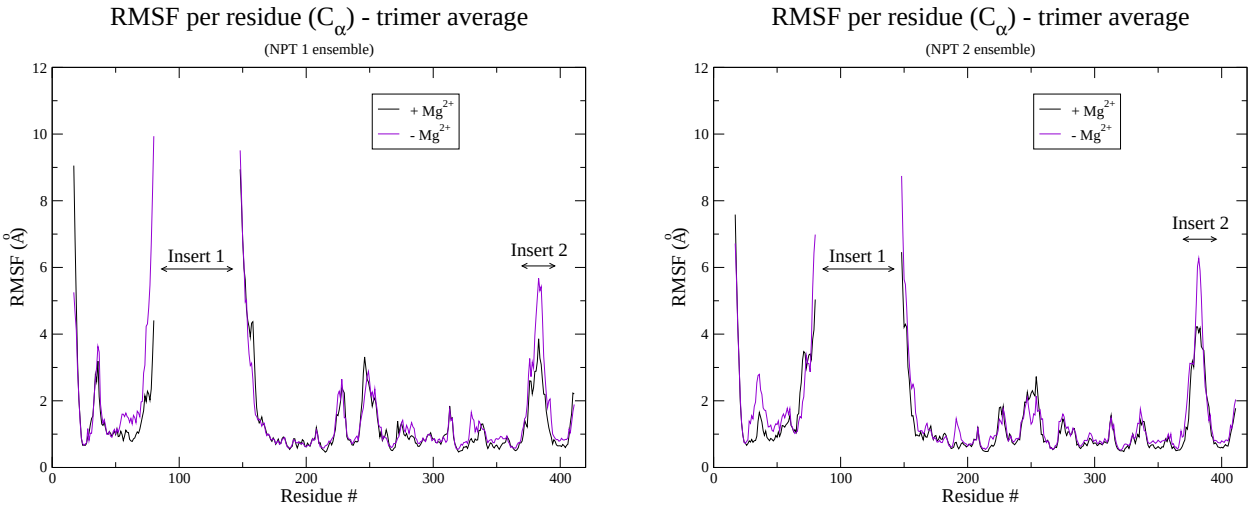


Figure 2.49: RMSF per residue (C_α atoms) for the NPT simulations. Compared to the NP simulation the same increase in movement of certain interacting regions without Mg^{2+} is not observed. Each chain was superimposed (all protein atoms) independently. The average for all three chains in each system (with and without Mg^{2+}) is shown. Every 100th frame was used (1000 frames total) for the calculation.

when measuring RMSF instead of RMSD certain regions that were observed to move more with Mg^{2+}removed in the NP ensemble could also be detected for NPT. Thus, RMSF seems to be a more sensitive measure for detecting differences between systems with and without Mg^{2+}. However, to obtain full agreement between NP and NPT it is expected that longer runs would be required. Due to time and hardware constraints this could not be attempted.

2.5 Conclusion

Increasing drug resistance to malaria highlights the need for new treatments. Lately, interfering with protein-protein interactions is being considered for potential drug discovery (Fry, 2006; Fletcher and Hamilton, 2006; Keskin *et al.*, 2007). *P. falciparum* arginase is notably different from the host equivalent in that enzyme activity and trimer formation are mutually

inclusive. In mammals it has been demonstrated that rat arginase I loses some activity (33–41% of *kcat*) when the trimerisation is disturbed by mutagenesis (Lavulo *et al.*, 2001). This has not been observed for human arginase I, where fully functional monomers have been obtained (Sabio *et al.*, 2001; Carvajal *et al.*, 1971, 1977). Despite the high sequence similarity (87%) between arginase I from rat and human (Di Costanzo *et al.*, 2007b), they differ in their kinetic properties. Human arginase I has a substantially lower K_m for arginine compared to rat arginase I. Furthermore, the K_d values for the inhibitors *S*-(2-boronoethyl)-L-cysteine and 2-amino-6-boronohexanoic acid are one order of magnitude less than for the rat counter part (Cavalli *et al.*, 1994; Baggio *et al.*, 1999; Kim *et al.*, 2001; Di Costanzo *et al.*, 2005). This suggests that it may be possible to inhibit *Pf*Arg via disturbing oligomerisation without affecting the human counterpart.

Arginine levels in *in vitro* cultures of *P. falciparum* are depleted by *Pf*Arg, although the relevance of arginase as a malaria drug target remains to be demonstrated (Olszewski *et al.*, 2009). Hypoargininaemia has been linked to the progression of severe malaria and may be related to the requirement of arginine for NO biosynthesis (Weinberg *et al.*, 2008). It has been speculated that low host NO benefits the parasite by causing increased expression of host intracellular adhesion molecule-1, which is used by parasitised red blood cells to adhere to the vascular endothelium and thus avoid spleen clearance. Arginase knockouts of the rodent malaria parasite, *P. berghei* (ANKA strain), are viable and show similar growth behaviour *ex vivo* and in infected mice (Olszewski *et al.*, 2009), although this has yet to be established for the human parasite.

Nonetheless, interfering with trimer formation in *Pf*Arg may serve as a parasite specific means of enzyme inhibition. To this end, *P. falciparum* arginase was modelled to further understand the structural metal dependency. Visual analysis of the model revealed a possible new inter-monomer salt-bridge between Glu 295 and Arg 404. Molecular dynamics simulations suggest that this interaction is susceptible to the removal of the active site metals. This was further corroborated by site-directed mutagenesis of the recombinantly expressed enzyme. Mutations expected to abolish the interaction promoted dissociation of the trimer and decreased enzyme activity. Simulations of *P. falciparum* arginase without the active site metals also resulted in a larger equilibrium C_α-RMSD, thus indicating that absence of the metals causes protein instability. However, the simulations were not long enough to observe dissociation or substantial protein unfolding. Longer simulations could therefore be pursued to further understand the possible linkage between active site metal binding and the Glu 295_a::Arg 404_b interaction. A further important result was that the general loss of structural integrity could be observed across different MD protocols, suggesting a real biological effect was observed and that this was not an artifact of the simulation. While some differences were observed between the NPT_1 and NPT_2 protocols, there was not sufficient time and resources to see if these could be replicated.

The results suggest that Glu 295_a::Arg 404_b interaction is necessary for trimer formation and enzyme activity of *P. falciparum* arginase, and forms part of the mechanism of

the structural metal dependency. The system used to simulate was incomplete in that it replaces Mn^{2+} with Mg^{2+} and does not completely simulate the co-ordination chemistry. Nonetheless, the results indicate this system could be used for *in silico* testing of molecules that interfere with trimer formation. Interfering with the formation of the Glu 295_a::Arg 404_b salt bridge may provide for novel inhibitors, since this interaction is not present in the host enzyme. The model system described could be used for *in silico* screening of small molecules designed to disturb the Glu 295_a::Arg 404_b and Arg 346_a::Glu 347_b interactions. Specifically, as Glu 347 Gln disturbs Glu 295_a::Arg 404_b suggests that a compound that induces similar behaviour may disrupt the enzyme.