Article

The Role of the Conformational Dynamics of Glutathione S-Transferase Epsilon Class on Insecticide Resistance in *Anopheles gambiae*

**Frederico J. S. Pontes,[a] Rafael T. Maia,[a,b] Maria Carolina P. Lima,[a]
Constância F. J. Ayres[b] and Thereza A. Soares*[,a]**

[a]*Departamento de Química Fundamental, Universidade Federal de Pernambuco,
Cidade Universitária, 50740-560 Recife-PE, Brazil*

[b]*Departamento de Entomologia, Centro de Pesquisas Aggeu Magalhães-FIOCRUZ,
Cidade Universitária, 50670-420 Recife-PE, Brazil*

Glutathione S-transferases (GSTs) are enzymes capable of metabolizing cytotoxic compounds. The enzyme AgGSTE2, member of epsilon class GSTs (GSTE), is the most important GST conferring resistance to dichloro-diphenyl-trichloroethane (DDT) in *Anopheles gambiae*. We have investigated the conformational dynamics of three GSTE variants (GSTE2, GSTE2-I114T/F120L, GSTE5) from *A. gambiae*. Large-scale motions of helices H2 and H4 and conformational transition of the C-terminal governs the opening of the G-site and is expected to affect substrate binding and product release. This structural rearrangement places Glu116 (Glu120 in GSTE5) close of the thiol group of the tripeptide glutathione (GSH) cofactor, making this residue a candidate to act as a base in the activation of DDT. The structural rearrangement is noticeable for AgGSTE2-F120L, which has been shown to confer increased DDT-resistance. The other variants exhibit a more subtle rearrangement. These findings corroborate the hypothesis that the increase of the conformational dynamics of GST Epsilon class isoforms from *A. gambiae* promotes higher DDTase activity.

Keywords: molecular dynamics simulation, evolutional constraint, positive and negative selection, metabolic resistance, malaria vector

Introduction

Glutathione transferases (EC 2.5.1.18) are highly promiscuous proteins where broad functional promiscuity co-exists with highly conserved structural fold. Glutathione S-transferases (GSTs) constitute a large family of cytosolic and membrane-bound proteins that catalyze the nucleophilic addition of the tripeptide glutathione (GSH) to a variety of electrophilic toxins and drugs (xenobiotics), leading to their excretion.[1,2] Lately, several other activities have also been associated with GSTs, including steroid and leukotriene biosynthesis, peroxide degradation, double-bond *cis-trans* isomerization, dehydroascorbate reduction, Michael addition, and noncatalytic ligand binding and transport activity.[3] These polymorphic enzymes belong to supergene families whose members are generated by punctual mutations, gene duplication and alternative splicing.[4] The GST family is subdivided into several classes accordingly to its occurrence in different taxa.

The number of isoforms *per* class varies widely, ranging from one to forty.[4] A single GST isoform is capable of conjugating glutathione to several hydrophobic substrates. Such promiscuity when coupled with the large number of isozymes generates a large range of potential substrates.

Insects exhibit at least six classes of GSTs (Sigma, Omega, Theta, Zeta, Delta and Epsilon) with the first four classes present in almost all living organisms.[4,5] The Epsilon and Delta classes are arthropod specific and some members of these classes have been associated to insecticide resistance in culicid vectors.[5,6] The metabolism of toxic compounds in insects is conducted by a series of enzymes from different phases and GSTs act in phase II. GSTs metabolize these compounds in two ways: one has been cited above (through GSH conjugation) and the other mechanism is by sequestration.[7] In conjugation reaction, the active site residue interacts with the GSH sulfydryl group generating the catalytically active thiolate anion (GS⁻). This nucleophilic thiolate anion attacks the electrophilic center of any lipophilic compound to form the corresponding GS⁻ conjugate.[8] The conjugation neutralizes the electrophilic

*e-mail: thereza.soares@ufpe.br

sites of the substrate, leading to its detoxification by the elimination of highly reactive electrophiles or rendering the product more water soluble and therefore more readily excretable from the cell.

Diversity of such enzymes is necessary to improve the way insects cope with environmentally harmful substances that imposes a limit on its niche occupation, particularly mosquito species that occupy a variety of habitats throughout its different stages of development. The promiscuity of GSTs is intimately connected to insecticide resistance, a multifactorial, multienzymatic and multigenic response to chemical toxic compounds used for insect control.[9,10] This biological phenomenon is observed on the individual and population levels, and has been a fundamental challenge for insect control programs due to the raise of resistant populations. One of the most known and reported insecticide resistances is observed for the *Anopheles gambiae* mosquito, the major malaria vector. Several studies have reported resistance to dichloro-diphenyl-trichloroethane (DDT) and other insecticides by this species.[11-13]

It has been shown that the GST Epsilon class isoform 2 from *A. gambiae* (AgGSTE2) is associated to DDT resistance via two mechanisms: elevated gene expression and catalytic activity of the recombinant enzyme.[14] For this reason, AgGSTE2 is overexpressed in resistant strains and the recombinant protein is efficient at metabolizing DDT.[14,15] Furthermore, we have previously sequenced six new GSTE members in three species of *Anopheles*, and shown that the GSTE2 gene displays the highest level of conservation among all the GSTEs in the four *Anopheles* species.[16]

This finding suggests a key role for GSTE2 in *Anopheles* adaptive processes whereas the other GSTE genes were more likely to undergo accelerated selection because of their relaxed evolutionary constraint. Indeed, we have further shown that one gene (GSTE5) is evolving under positive selection.[16] Recent work has associated DDT resistance in *A. gambiae* with the occurrence of target-site resistance mechanisms introgressed between emerging species, and the appearance of allelic variants in GSTE2.[1] It has also reported on the sequencing of GSTE2 from DDT resistant and susceptible strains of *A. gambiae*, showing the incidence of polymorphism (I114T) that segregated with strain phenotype. Furthermore, kinetic assays have shown that at high concentrations DDT inhibits GSTE2 from the susceptible strain whereas DDT is converted into its product DDE (1,1-dichloro-2,2-bis(*p*-chlorophenyl) ethylene) by GSTE2 expressing the I114T substitution.[1] Notwithstanding, the X-ray structures of GSTE2 and GSTE2-I114T do not exhibit significant structural

differences that could explain their distinct catalytic efficiencies.[1]

It has been postulated that the increased catalytic activity of the GSTE2-I114T variant, shown to impart DDT resistance to *A. gambiae*, result from increased conformational dynamics with respect to the GSTE2 isoform.[1] Increased conformational dynamics has also been proposed to explain the 350-fold higher catalytic efficiency of GSTE2 compared to GSTs from the Delta class in the same species.[14,17] Previous structural and biochemical characterization of the Delta class GSTD4 from *A. gambiae* has underlined the importance of several residues for catalysis and substrate specificity of the enzyme through structural rearrangements.[18,19] These conformational changes have been postulated to alter the flexibility of the protein, and thus modulate its catalytic aptness.[1]

The enzyme AgGSTE2 is thought to be the most important GST conferring DDT resistance in *A. gambiae*. It is a promising molecular target for the enhancement or design of insecticides against severe insect transmitted diseases such as malaria, yellow fever and dengue fever. In this report, we investigate the proposed role of increased conformational dynamics on the biological activity of GSTE variants (GSTE2, AgGSTE2-I114T/F120L and GSTE5) from *A. gambiae* via molecular dynamics (MD) simulations (Figure 1).[1,17] MD simulations provide an atomic description of the time-evolution of conformations of a system due to its singular property of probing space and time scales simultaneously.[20-22] For this
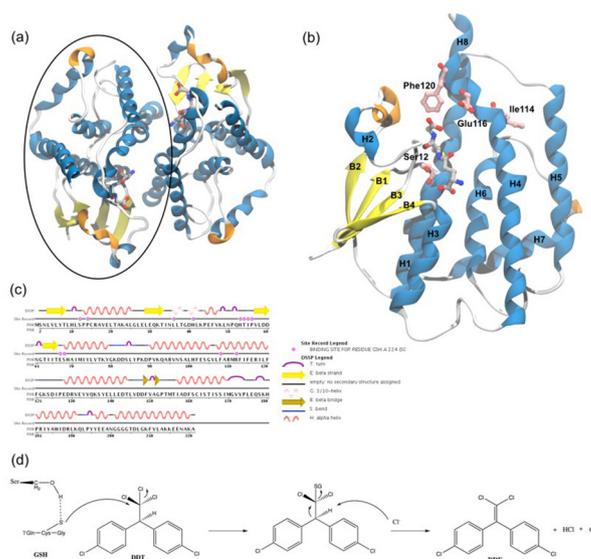

Figure 1. Cartoon representation of the X-ray structure of AgGSTE2 (PDB ID 2IMI). (a) Homodimer; (b) active site in one monomer. Secondary-structure elements are labeled according to Wang *et al.*;[17] (c) secondary structure assignment for amino acid sequence from the Protein Databank (www.rcsb.org). Only the monomer sequence is represented in panel C. Residues Ser12 and Glu116 correspond to residues Ser15 and Glu120 respectively in AgGSTE5; (d) schematic representation of the catalysis of DDT by AgGST.

reason, it is a powerful technique to explore conformational transitions of biomolecules at the atomic level. We report on the structural rearrangements of the DDT pocket cap, which led to the partial occlusion of the G-site, and displacement of the Glu116 (Glu120 in GSTE5) towards the thiol group of GSH upon GSH binding. The conformational change of the DDT pocket cap is expected to affect substrate binding and/or product release, influencing the catalytic turnover of the enzyme. Further, the atomic displacements underlying this motion are noticeably higher for AgGSTE2-F120L, which has been shown to confer increased DDT-resistance in *A. gambiae*. The GSTE2 and GSTE5 variants, which confer low or no DDT-resistance, exhibit lower amplitude displacements. These findings support the hypothesis that the increase of the conformational dynamics promotes higher DDTase activity of GST Epsilon class isoforms from *A. gambiae*.[1,17] We propose a molecular mechanism to explain the effect of GST polymorphisms on DDT-resistance.

Experimental

Methodology

Comparative structural modeling of the AgGSTE5 sequence

Comparative protein structure modeling technique was applied to build a structural model of the target sequence of AgGSTE5 on the basis of sequence alignment with the template X-ray structure of AgGSTE2.[17] Protein structure homology modeling relies on the evolutionary relationship between the target and template proteins.[23,24] Therefore, the more similar two sequences are, the closer the corresponding structures can be expected to be and the larger the fraction of the model that can be directly inferred from the template. The GSTE5 primary sequence was obtained from VectorBase (accession number: AGAP009192-PA). The X-ray structure of AgGSTE2 (E.C. 2.5.1.18) from *Anopheles gambiae* solved at 1.4 Å resolution is available in the PDB database (PDB IDs 2IL3 and 2IMI). The X-ray atomic coordinates were used as template to build the structural model of the target sequence corresponding to the AgGSTE2 sequence from the same species. Pairwise alignment showing high

target-template similarity were retrieved by a BLAST search against the SWISS-PROT database via the SWISS-MODEL web interface.[25,26] The BLOSUM62 substitution matrix was used with gap open and gap extend penalties of 11 and 1 respectively.[27] It has been shown that sequence alignments can unambiguously distinguish between protein pairs of similar and non-similar structures when the pairwise sequence identity is higher than 30%.[28,29] Accordingly, the BLAST alignment between target and template sequences displayed a significant sequence homology with similarity and identity levels of 76 and 52%, respectively (Supplementary Information section, Figure S1).

As a rule, regions of the target protein aligned to the template can be modeled on the basis of the information provided by template structures,[29,30] whereas gap regions that are not aligned with a template (insertions/deletions) require specialized approaches.[31-35] The SWISS-MODEL workspace has been used to construct the structural models of the target sequences.[36] Template structures are superposed by the means of a least squares algorithm and a structural alignment is generated after removing template structures with high C root mean square deviations to the highest homology template. A local pair-wise alignment of the target sequence to the main template structure is calculated, followed by attempts to improve the alignment with the other template structures for modeling purposes. Sequence alignment exhibited only two gaps of one amino acid each out of a total of 218 amino acids (Supplementary Information section, Figure S1). Regions of amino acids insertions or deletions in the target sequence (gaps) were modeled via the SWISS-MODEL scoring scheme.[36] In this approach an ensemble of fragments compatible with the neighboring stems is constructed by selecting the best loop which accounts for force-field energy, steric hindrance and favorable interactions like hydrogen bond formation.[37] If no suitable loop can be identified, the flanking residues are included to the rebuilt fragment to allow for more flexibility. In cases where the procedure cannot produce a satisfying solution and for loops above 10 residues, a loop library derived from experimental structures is searched to find compatible loop fragments. The ANOLEA mean

Table 1. Simulated systems

	System	PDB ID	Solute atom	Ion	Water molecule
Apo enzyme	AgGSTE2	2il3	4414	165	73986
	AgGSTE2-I114T/F120L	2il3	4417	150	68325
	AgGSTe5	2il3	4456	158	68352
Holo enzyme	AgGSTE2	2imi	4487	164	73632
	AgGSTE2-I114T/F120L	2imi	4471	164	73653
	AgGSTe5	2imi	4510	172	73737

force potential were used to evaluate the local quality of the structural models.[38] The stereochemical plausibility of the models was evaluated via comparisons of Ramachandran plots calculated for experimental and modeled data (Supplementary Information section, Figure S2).

The reconstruction of the model side chains was based on the weighted positions of corresponding residues in the template structures.[26,36] The model side chains were built by iso-sterically replacing the template structure side chains with side chain conformations selected from a backbone dependent rotamer library.[39] A scoring function assessing favorable interactions (hydrogen bonds, disulfide bridges) and unfavorably close contacts was applied to select the most likely conformation.[26,36] Deviations in the protein structural geometry may be introduced when joining rigid fragments in regions of low sequence homology with the template structures. In such instance, molecular dynamics simulations are necessary to improve side-chain packing and favorable interactions. Explicit-solvent molecular dynamics simulations were carried out for this system and are described in the molecular dynamics simulation subsection. The AgGSTE5 model structure was deposited in the Protein Model Data Base.

Molecular docking of DDT to AgGSTE variants

Crystallographic structures of AgGSTE2 (PDB ID 2IMI[1] and 4GSN[17]) from *A. gambiae* were used as initial coordinates for the molecular docking calculations. Atomic coordinates for AgGSTE2-I114T/F120L were generated by replacing the side chain of residues I114 and F120 by threonine and leucine, respectively, in the X-ray structure 2IMI.[1] Protonation states were assigned accordingly to pH 7. Molecular docking simulations were performed with the program Autodock4[40-42] and AutoDock Tools.[43] The partial charges on the protein atoms were taken from the AMBER all-atom force field.[44] The grid maps were calculated using AutoGrid.[45] A grid map with $129 \times 129 \times 129$ points and a grid-point spacing of 0.25 Å was centered on the atomic coordinates of the GSH cofactor. During the docking simulations, dihedral angles were treated as flexible for DDT. The Lamarckian genetic algorithm as implemented in AutoDock4 program[40-42] was used with the following parameters: an initial population of 100 random individuals, a maximum number of 1.5×10^6 energy evaluations, a maximum number of 27000 generations with mutation and crossover rates of 0.02 and 0.08, respectively. An optional elitism parameter equal to 1 was applied. It determines the number of top individuals that will survive into the next generation. A maximum of 300 iterations *per* local search was allowed. The probability of performing a local search on an individual was 0.06 where the maximum number of

consecutive successes or failures before doubling or halving the search step was 4. After the conformational search, the docked conformations were sorted in order of increasing energy. The coordinates of the lowest energy conformations were clustered based on a root-mean-squared-deviation (RMSD) of 2.0 Å. A more detailed description of the methodology employed here can be found elsewhere.[46,47]

Setup and molecular dynamics simulations

MD simulations were performed for glutathione S-transferase isoforms AgGSTE5 and AgGSTE2 from *Anopheles gambiae* (Table 1). High-resolution crystallographic structures of AgGSTE2 from *A. gambiae* were used as initial coordinates for simulations of the apo (PDB ID 2IL3) and the holoenzyme (PDB ID 2IMI) forms.[17] The comparative structural model was used for simulations of both apo and holoenzyme AgGSTE5 (Table 1). Two single-mutations (I114T, F120L) of the AgGSTE2 associated to DDT-resistance in *A. gambiae* populations were also modeled. Hereafter, the wild type and double mutant GSTE2 proteins will be referred to AgGSTE2 and AgGSTE2-I114T/F120L. Protonation states were assigned accordingly to pH 7, except for histidines residues spatially clustered together. In this case, multiple MD simulations were performed to assess the combination of protonation states compatible with the preservation of the dimer arrangement in the dimer. Therefore, the protonation state of residues His10, His53 and His69 corresponded to a charge of +1 in the holoenzymes whereas only residue His69 had a charge of +1 in the apoenzymes. The remaining histidines residues were kept at a protonation state with charge 0. Simulations were carried out using the GROMOS force field force parameter set 54A7[48-50] in conjunction with the software package GROMACS v.4.5.4 double precision.[51,52] Atomic parameters for DDT and GSH were taken from GROMOS force field force parameter set 54A7. The topologies and assigned atomic parameters for the two compounds are presented in the Supplementary Information section (Figures S3 and S4).

The systems were placed in a rectangular box of dimensions $9.5 \times 8.5 \times 9.5$ nm³, treated for periodic boundary conditions, and solvated with explicit SPC model water molecules.[53] The total charge of the systems was neutralized by adding Na^+ counterions, and subsequently Na^+ and Cl^- counterions were added to yield an ionic strength of 150 mmol L⁻¹ (see Table 1). Simulations were carried out in the constant-temperature, constant-pressure (NPT) ensemble, and a time step of 1 fs were used to integrate the equations of motion based on the Leap Frog algorithm.[54] The temperature of the solute and solvent were separately coupled to the

Nosé-Hoover thermostat at 300 K with a relaxation time of 0.2 ps.[55,56] The pressure was maintained at 1 atm by isotropic coordinate scaling via the Parrinelo-Rahman barostat with a relaxation of 1 ps.[57] Bond lengths were constrained via the P-LINCS algorithm,[58] and the geometry of the water molecules was constrained using the SETTLE algorithm.[59] The Particle-Mesh Ewald correction with a maximum grid spacing of 0.12 nm and a cubic interpolation order.[60,61] A cutoff of 1.4 nm was used for both vdW and long-range electrostatic interactions. In all cases, the pair list for short-range nonbonded and long-range electrostatic interactions were updated with a frequency of 5 timesteps. The center of mass motion was removed at every 5 steps. The systems were initially minimized through 2000 iterations using the steepest descent algorithm. Solvent molecules were relaxed during 500 ps at 300 K with positional restraints applied to the backbone atoms of the protein. The temperature was slowly increased at intervals of 50 K whereas distance restraints between residue pairs Asp129-Arg130, Phe113-Arg130 (AgGSTE2), Phe117-Phe340 (AgGSTE5) were applied during the 5 ns period of equilibration period. Production phase was carried out for 50 ns. Configurations of the system were recorded as trajectory files every 1 ps. Protein structures were visualized with the software VMD 1.86.[62] Structural properties were computed using standard implementation of algorithms in the GROMACS package.[51,52]

Results and Discussion

DDT-binding to AgGSTE variants

We have investigated the binding mode and affinity of DDT for AgGSTE2 variants containing either Phe or Leu at position 120 and for AgGSTE5 via molecular docking calculations. During the calculations, at least 4.05×10^8

conformations of DDT were sampled *per* molecular docking calculation for each GST variant. The DDT molecule binds to AgGSTE2 in a well-defined conformation, which places the β-hydrogen of the ligand within ca. 2 Å of the thiolate group of GSH. This conformation is consistent with an elimination reaction for the conversion of DDT into DDE as proposed from the manual modeling of DDT into the active site of the X-ray structure of AgGSTE2.[17] DDT adopts the same conformation upon binding to the AgGSTE2 variants, regardless of the presence of Phe120 or Leu120 (Figure 2). However, alternate conformations of Phe120 in the X-ray structures of AgGSTE2 (PDB ID 2IMI and 4GSN) can either hinder or assist the correct placement of DDT with respect to GSH during the calculations (Figure 2).[1,17] For instance, Phe120 conformation in crystallographic structure 4GSN places the β-hydrogen of DDT within ca. 4.5 Å of the thiolate group of GSH, thus rendering the conformation unsuitable for catalysis.

We have performed MD simulations of AgGSTE2 bound to DDT to assess the role of Phe120 during the inhibitor binding. Besides Phe120, residues Phe115, Phe121 and Phe210 delimit the bottom of the G-site pocket. DDT binds to the G-site through hydrophobic contacts with the four aromatic residues. These residues place the β-hydrogen of DDT towards the thiolate group of GSH for catalysis (Figure 3). The MD simulations show that Phe120 is the main anchorage point for DDT via extensive π-stacking interactions between their respective aromatic rings. Phe120 adjusts conformationally to such task, being helped by Phe121. Hence, Phe120 appears to facilitate the binding of DDT to AgGSTE2, and steric hindrance cannot explain the comparatively higher binding affinity of DDT for AgGSTE2 containing Leu (50.9 µmol L⁻¹) *versus* Phe (97.8 µmol L⁻¹) at position 120. We hypothesize that large-scale conformational rearrangements of secondary structure
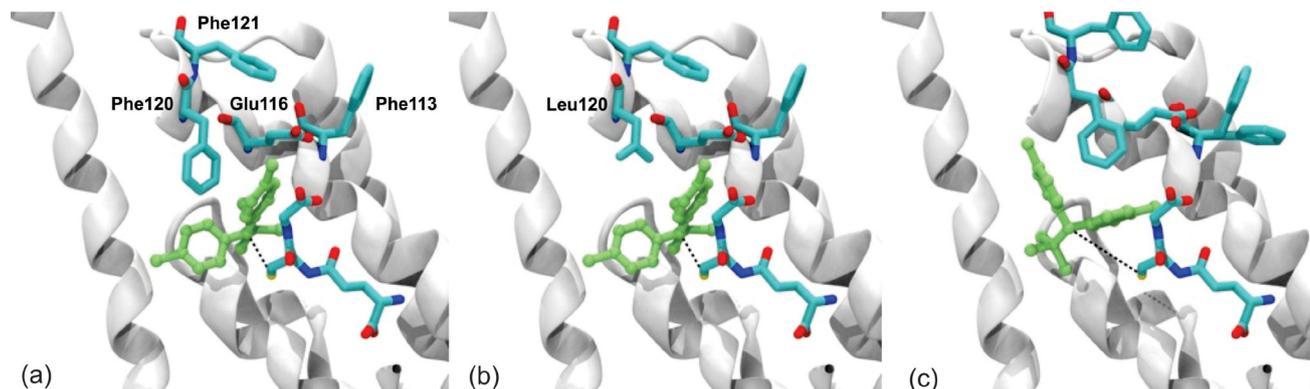

Figure 2. Predicted binding conformation of DDT to AgGSTE2 variants. (a,b) Atomic coordinates from X-ray structure 2IMI solved at 1.4 Å resolution; (c) atomic coordinates from X-ray structure 4GSN solved at 2.4 Å resolution.[1,17] AgGSTE2-I114T/F120L was built by replacement of side chain atoms in mutated residues. DDT is shown in green ball-and-stick. Residues are represented in stick with carbon atoms in cyan, nitrogen in blue, oxygen in red, sulfur in yellow.

elements may underlie differences in binding affinity of AgGSTE2 to DDT. This is discussed in the next section.

Molecular docking of DDT into the active site of AgGSTE5 did not produce any GSH-bound conformer. Visual inspection of the structural model of AgGSTE5 shows that the region around the thiolate group of GSH is too tight for DDT to bind. This may result from the structural inaccuracies inherent to the homology model or it may indicate that AgGSTE5 lacks DDTase activity. In fact, it was shown that AgGSTE1, AgGSTE3 and AgGSTE4, and most likely AgGSTE6 and AgGSTE7, do not exhibit DDTase activity.[14] The apparent inaptness of AgGSTE5 to bind DDT may suggest that this isoform does not exhibit DDTase activity as well. However, such hypothesis needs corroboration through enzyme kinetic measurements.

Structural dynamics of the dimer subunits

Atom positional root-mean-square deviation (RMSD) was calculated from the MD simulations of the apo and holo forms of AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5 with respect to the respective X-ray structures (PDB ID 2IL3 and 2IMI).[17] The RMSD profiles for each subunit are similar among the three apoenzymes, and to a smaller degree, among the holoenzymes (Figure 4). The time-dependent secondary structure content for the simulated systems exhibits the same pattern (Supplementary Information section, Figure S5). This is consistent with a well-conserved subunit fold throughout the simulations. Conversely, the RMSD profiles calculated from the superposition of atoms in both subunits onto the X-ray coordinates are fairly higher and do not reach convergence within the 50 ns time scale. The behavior is more noticeable for the holoenzymes (i.e., upon GSH binding), and results from the large amplitude motion of the dimer subunits with respect to each other as discussed in the next section.

Conversely, the RMSD profiles calculated from the superposition of atoms in both subunits onto the X-ray coordinates are fairly higher and do not reach convergence within the 50 ns time scale. The behavior is more noticeable for the holoenzymes (i.e., upon GSH binding), and results from the large amplitude motion of the dimer subunits with respect to each other as discussed in the next section.

The atom positional root-mean-square fluctuation (RMSF) profiles obtained from the MD simulations show that regions of higher flexibility for AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5 are similar, but the amplitude of the atomic fluctuations can differ between the apo and holoforms and between subunits of the same protein (Figure 5). The most remarkable aspect of the RMSF profiles is the large atomic fluctuations of the C-terminal region. The increased flexibility of the C-terminus has been previously suggested based on the X-ray structure of AgGSTE2.[17] A highly flexible C-terminal region is also characteristic of GST members of the Alpha class in humans.[11] In this class, structural homologs exhibit distinct promiscuity profiles, which are correlated with either a well-defined (non-promiscuous enzyme) or a highly disordered (promiscuous enzyme) conformation of the C-terminal helix. In the promiscuous human GSTa1, residues 212-222 in the C-terminal helix are highly disordered and cannot be defined in the X-ray structure of the apo-enzyme.
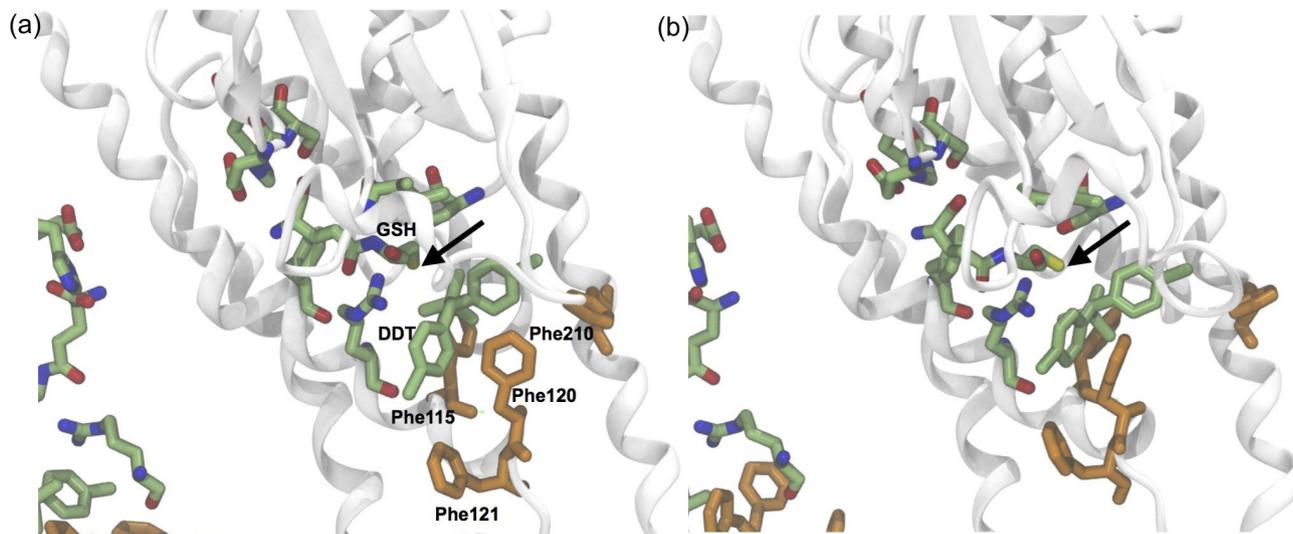

Figure 3. (a) Initial; (b) final conformation of AgGSTE2 bound to DDT. Residues are represented in stick with carbon atoms in green, nitrogen in blue, oxygen in red, sulfur in yellow. Phenylalanine residues surrounding DDT are shown in orange. Black arrows indicate the thiolate anion from GSH. Hydrogen atoms are not shown for clarity.

In the AgGSTE2 and AgGSTE2-I114T/F120L simulations, the C-terminal region of one of the subunits folds into a β-strand to compose an antiparallel β-sheet

with the loop H4-H5 (Figure 6). The β-sheet precludes the helix H4 from bending over the G-site, leaving GSH less constrained conformationally. The β-sheet does not
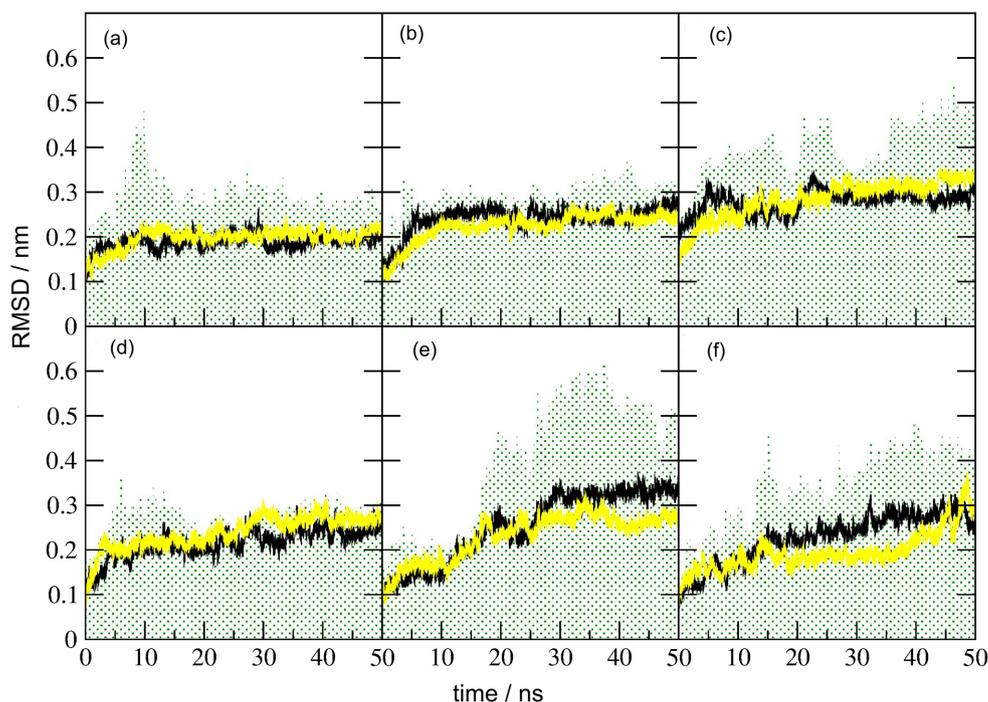

Figure 4. Root-mean-square deviation of Cα atoms from X-ray structure for the apoenzymes (a) AgGSTE2; (b) AgGSTE2-I114T/F120L; (c) AgGSTE5; and the holoenzymes (d) AgGSTE2; (e) AgGSTE2-I114T/F120L; (f) AgGSTE5. Root-mean-square deviations were calculated for Cα atoms in subunit A (black), subunit B (yellow) and in the full dimer (green dots). The X-ray structures 2IL3 and 2IMI were used as positional reference for the apo and holoenzymes, respectively.
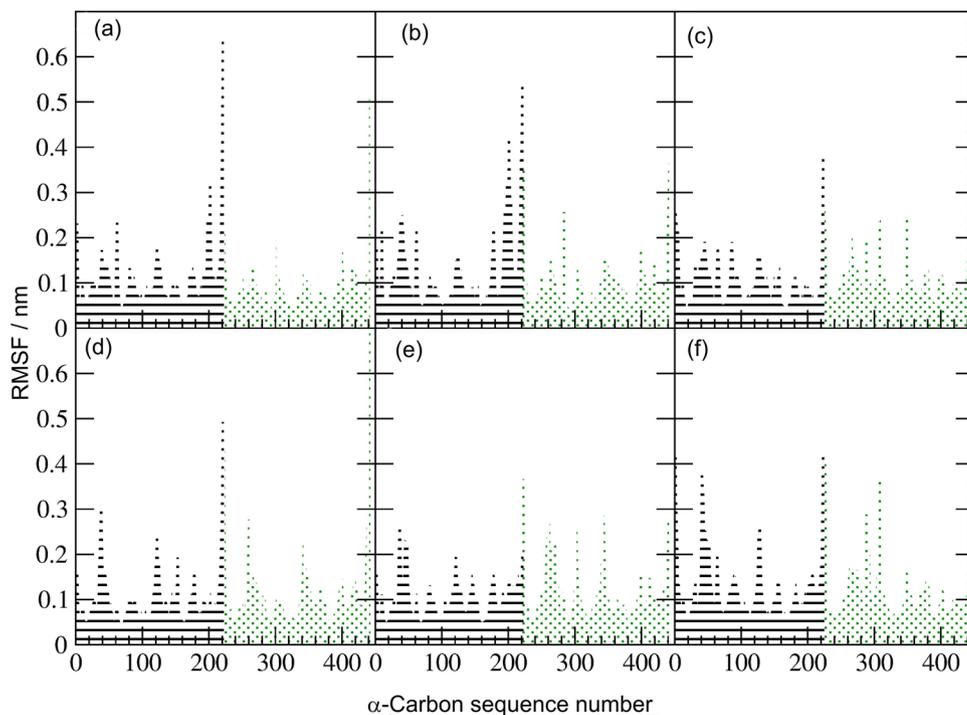

Figure 5. Root-mean-square fluctuation of Cα atoms for the apoenzymes (a) AgGSTE2; (b) AgGSTE2-I114T/F120L; (c) AgGSTE5; and the holoenzymes (d) AgGSTE2; (e) AgGSTE2-I114T/F120L; (f) AgGSTE5. Rotational and translational fitting of pairs of structures was applied using Cα atoms. RMSF values were calculated for the final 25 ns of simulation.

occur in the simulations of AgGSTE5 within the time scale of 50 ns. Other regions of increased flexibility comprise residues 35-45 (loop B2-H2), 111-124 (H4) 216-221 (H8) in both subunits, and residues 60-68 (B3-B4) and 80-86 (loop H3-H4) (Figure 5). Regions of increased atomic fluctuations differ between the two subunits of the same protein, suggestive of asynchronous motions between the dimer subunits upon GSH binding. Overall, the RMSD and RMSF profiles are equivalent between the apo and corresponding holoenzymes. The present simulations of apoenzymes were intended to serve as reference for comparison with the biologically relevant holoform. For this reason, the discussion hereafter will focus on the holoforms.

Large-scale conformational rearrangements of the GSTe isoforms

The technique of principal component analysis of MD trajectories was used to separate internal motions into orthogonal motions.[63] Large-amplitude motions of AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5 are largely described by the first eigenvector mode as shown by its associated eigenvalues (Figure 7).

No significant large vibrational motion was observed along the remaining eigenvectors for AgGSTE2 and AgGSTE5. The atomic displacement described by the first eigenvector has eigenvalues of ca. 8, 3 and 2 nm^2 for AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5, respectively (Figure 7A). The contributions of the backbone atoms to the first eigenvector represent the relative displacement of each residue due to the motion described by that eigenvector (Figure 7B). For AgGSTE2 and AgGSTE2-I114T/F120L, the movement along the first eigenvector is mainly concentrated in the helix H5 (residues 110-120), in the short helix H2 (residues 35-50), which connects two antiparallel strands of the β-sheet, and in the C-terminal region (Figure 6).

Some contribution from residues 200-210 in the C-terminal helix is also observed, with all three regions located around the GSH binding site (Figure 1). Likewise, the regions of increased atomic displacement in AgGSTE5 comprised the C-terminal helix (residues 210-216), the loop connecting helix H2 (residues 50-54) and helix H5 (residues 110-120) (Figure 7B). However, the contributions of distinct structural regions to the first eigenmode are more homogeneous (and less accentuated) for AgGSTE5 than for the remaining systems (Figure 7B). In brief, equivalent structural regions contribute for the largest scale motions in AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5 albeit with considerably different amplitudes.

The variant AgGSTE2-I114T/F120L exhibits the largest amplitude displacements of the C-terminal, H2 and H5 helices whereas AgGSTE5 exhibits the smallest amplitude displacements for the same regions.

The conformational transitions associated with large amplitude motions occur through correlated motions (Figures 7C-E). First, residues helix H5 (residues 110-120) and C-terminal helix H9 (residues 200-210) move towards each other to produce a breathing movement between the upper portions of the two subunits (Figure 7C). Then, helix H2 (residues 35-50) slides down towards the β-sheet, leading to the partial occlusion of the G-site. In AgGSTE2 and AgGSTE2-I114T/F120L, this sequence of events takes place in only one of the subunits (within 50 ns of simulation) (Figures 6c-d). In the other subunit, the rearrangement of the C-terminal terminus into an antiparallel β-sheet with the C-terminus of helix H4 induces the opening of the G-site (Figures 7C-D).

In the AgGSTE5, the breathing movement involving the upper portions of the two subunits is also observed, but at much smaller amplitude as anticipated from analysis of the corresponding eigenvalues (Figure 7B). For this protein, the largest amplitude motion is dominated by the atomic displacement of the C-terminus.

Interactions of residues in the G-site upon GSH binding

The MD simulations of AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5 unveil structural rearrangements of large-amplitude involving three major regions: loop S2-H2 connecting strand S2 and helix H2, the C-terminus of helix H4 (residues 120-124) and the C-terminal region (Figure 1). These rearrangements lead to the occlusion of the G-binding site when in presence of GSH, which participates through interactions to residues in the loop S2-H2 and C-terminus of H4. The movement is asynchronous between the two subunits of the dimer in AgGSTE2 and AgGSTE2-I114T/F120L within the 50 ns timescale. Two distinct conformations can be discerned in the MD-derived structural ensembles. In the open conformation, the loop S2-H2 is pulled away from the G-binding site, the C-terminal makes an anti-parallel beta-strand with the C-terminus of the helix H4 (except for AgGSTE5), and GSH is loosely bound to the G-binding site as shown by high atomic fluctuations and diffusion coefficients (Figure 6). It should be noticed that the beta-strand is transient in AgGSTE2 but not in AgGSTE2-I114T/F120L where it occurs during the full 50 ns of simulations. In the closed conformation, the helix H4 bends and places its C-terminal region on the top of the G-binding site. It occurs concomitant with the displacement of the loop
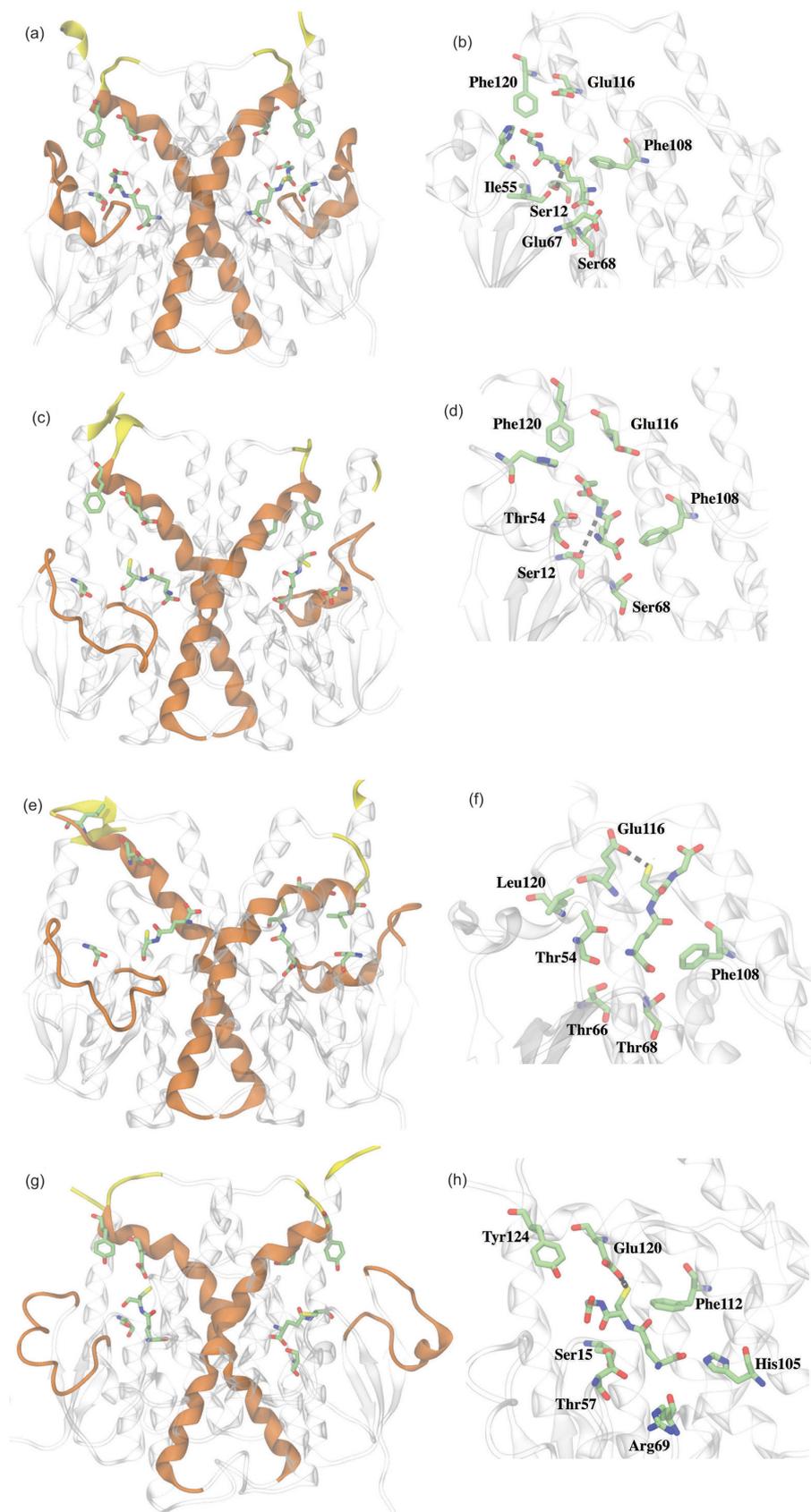
Figure 6. Conformational changes involving helices H2, H4 and the C-terminal region of AgGSTE variants during MD simulations. The structural rearrangement places Glu116 within hydrogen bond distance of the thiol group of GSH. (a,b) X-ray structure (2IMI); (c,d) final conformations from MD simulations of AgGSTE2; (e,f) AgGSTE2-I114T/F120L; (g,h) AgGSTE5.
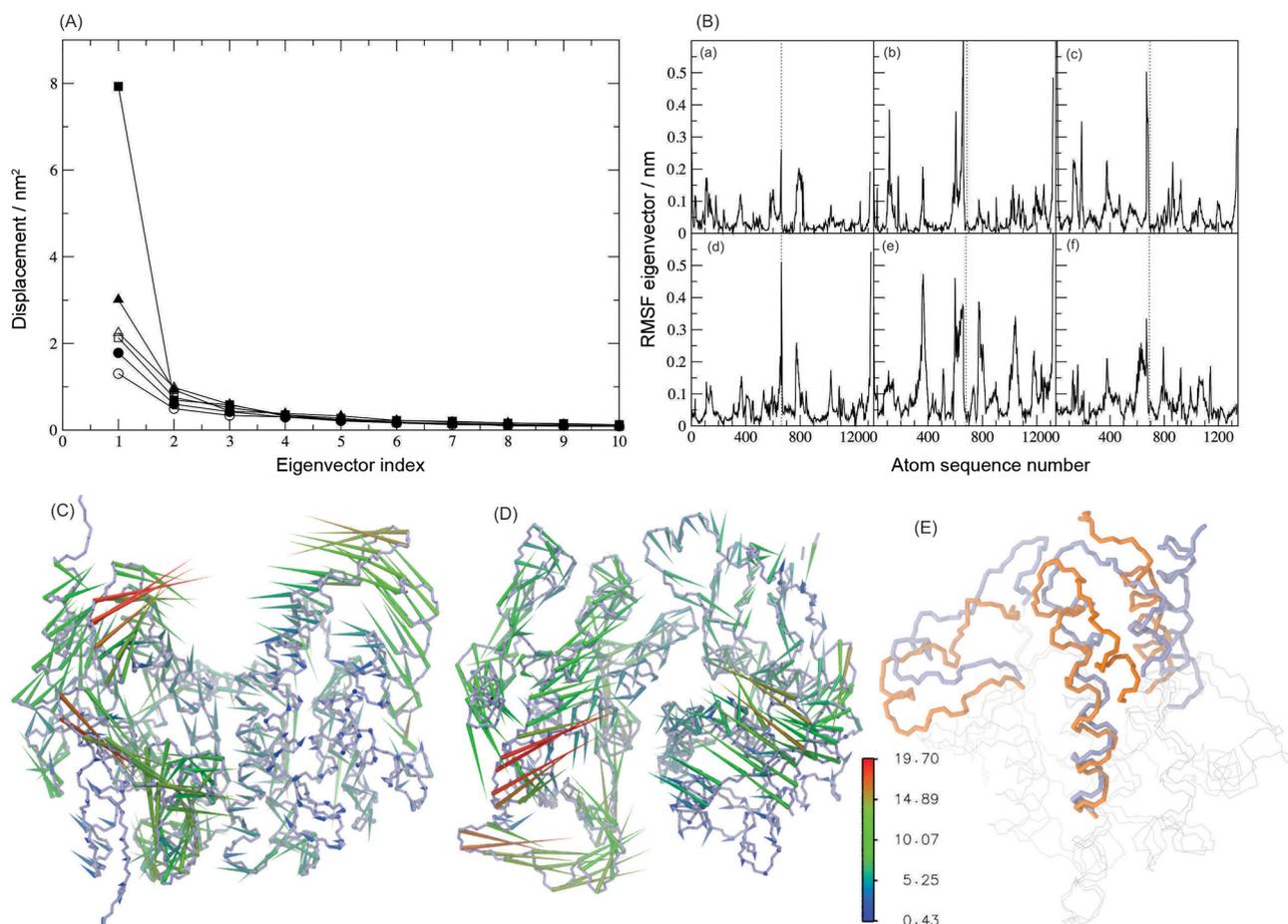
Figure 7. Principal component analysis of MD trajectories. (A) Eigenvalue amplitude for systems AgGSTE2 (circle), AgGSTE2-I114T/F120L (square) AgGSTE5 (triangle). Apoprotein and holoprotein are represented by open and closed symbols, respectively. (B) Representation of Ca atomic displacement along the first eigenvector for apoenzymes (a) AgGSTE2; (b) AgGSTE2-I114T/F120L; (c) AgGSTE5, and holoenzymes (d) AgGSTE2; (e) AgGSTE2-I114T/F120L; (f) AgGSTE5. (C,D) Porcupine plots representing the highest amplitude motions along the first eigenvector. (E) Superposition of Cα atomic coordinates corresponding to the maximum and minimum projections of the first eigenvector for AgGSTE2-I114T/F120L-GSH.

S2-H2 towards the G-binding site and the C-terminus of helix H4. These structural rearrangements are observed to a larger or minor degree in the MD simulations of AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5. However, the structural transition leading to the closed conformation is triggered by local interactions that differ among the three proteins. In the AgGSTE2 simulation, hydrogen bonds between His41 and the glycyl carboxylate in GSH pulls the loop S2-H2 towards the G-binding site. Subsequently, the C-terminal Lys220 interacts via hydrogen bonds with the carbonyl groups of residues Leu37, Thr38 and Gly39 in the loop S2-H2. This movement disrupts the transient beta-strand between the protein C-terminus and the C-terminus of helix H4, and further stabilizes the placement of the loop S2-H2 over the entry of the G-binding site. In the closed conformation, GSH binds tightly to residues Ser12, His41, Gln52, Thr54, Ser68 and Arg112. In the AgGSTE2-I114T/F120L simulation, there is the formation of a hydrophobic nucleus involving Leu37 and residues in the C-terminus of

helix H4 (Leu119, Leu120 and Phe121), which prompts the motion of the loop S2-H2 towards the G-binding site. These interactions are facilitated by the pronounced bend of helix H4, particularly when compared to AgGSTE2 and AgGSTE5. The rearrangement places GSH to bind residues Thr54, Ser68, Arg112 and Asp116, and transiently with residues Gln52 and Glu67. The displacement of loop S2-H2 and the C-terminus of helix H4 in AgGSTE5 also occurs via the establishment of a hydrophobic nucleus between residues in two regions. The residues involved in the displacement are Leu123 and Tyr124 (helix H4) and Leu39 and Leu40 (loop LS2-H2).

Potential role of Glu116 in the activation of GSH

Previous mechanistic investigations of GSTs have underlined the importance of a conserved tyrosine or serine residue at the glutathione binding site in facilitating the deprotonation of GST thiol group.[19,64-67] The catalytic

activity of GSTs relies on the enzyme ability to activate GSH by lowering the pKa of its thiol group from 9.0 to values between 6.0 and 6.9.[65] It enhances the rate of the nucleophilic attack of GSH towards electrophilic co-substrates in 200 to 300-fold. Two classes of GSTs can be recognized based on the identity of the residue (Tyr or Ser) acting as a base during the activation of GSH. However, there has also been experimental support for an alternative mechanism. In the base-assisted deprotonation model, the glutamyl α-carboxylate of GSH acts as the catalytic base deprotonating the thiol group.[65,68] A variation of this model has been proposed for the catalytic mechanism of Delta class GSTs.[18,19] In the model known as base-assisted deprotonation the glutamyl α-carboxylate of GSH acts as the catalytic base deprotonating the thiol group.[65,68] On the basis of the structural similarity and evolutionary proximity between the Delta and Epsilon classes, the proposed mechanism is also plausible for AgGSTE2 and AgGSTE5.[69]

In the X-ray structure of AgGSTE2, the thiol group of GSH is within hydrogen-bond distance of the hydroxyl group of Ser12, which is thought to be required for the correct orientation and stabilization of the deprotonated thiolate anion in the active site (Figure 1).[19,65,66] Early in the simulations of AgGSTE2-I114T/F120L and AgGSTE5, the thiol group moves away from Ser12 and

approaches the carboxylate group of Glu116 (Figure 8). This conformational change is more noticeable in the subunit where the S2-H2 loop and H4 helix shut the G-site as opposed to the subunit where the G-site remains more exposed to the solvent (Figure 7). In the AgGSTE2-I114T/F120L and AgGSTE5 simulations, the conformational changes associated with the interaction between the thiol group and Ser12 or Glu116 occurs through sharp transitions. In the AgGSTE2 simulation, Glu116 approaches the GSH thiol group (e.g., at 15 and 35 ns) but the distance between the two groups remains overly large to allow proton withdraw. The distances between GSH-Ser12 and GSH-Glu116 are anti-correlated in the three simulations.

Since Glu116 is expected to have a much lower pKa than Ser12, it could be a more likely candidate to act as a base in the activation of GSH. In this scenario, Glu116 could deprotonate the thiol group of GSH, and the negatively charged GS⁻ would be stabilized by interactions with the hydroxyl group of Ser12 and other nearby residues. Based on the conformational dynamics of the three variants, the activation of GSH by Glu116 either does not occur for AgGSTE2 or occur at lower frequencies for AgGSTE2 than for AgGSTE2-I114T/F120L and AgGSTE5. As result, there would be a larger fraction of activated GSH available for DDT-binding in the AgGSTE2-I114T/F120L and AgGSTE5
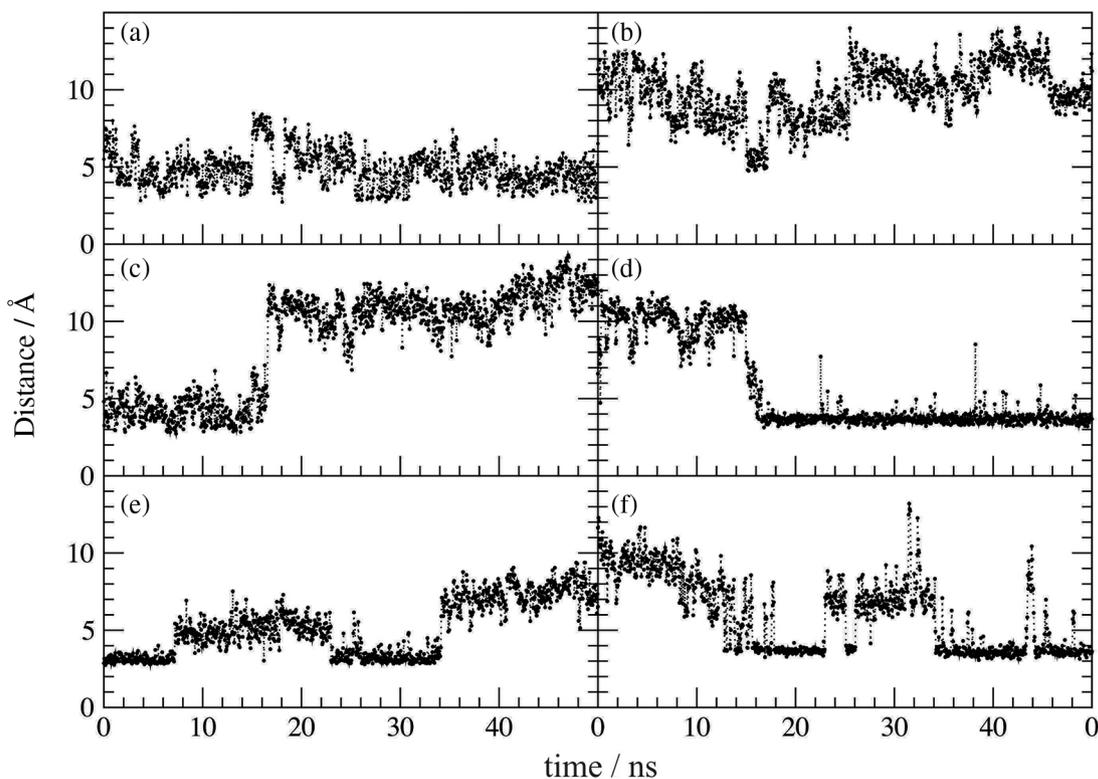

Figure 8. Time-dependent distances between Ser12-GSH (a, c, e) and Glu116-GSH (b, d, f) for (a, b) AgGSTE2, (c, d) AgGSTE2-I114T/F120L, and AgGSTE5 (e, f). Distances were calculated between the center of mass of the hydroxyl group of Ser12 or carboxylate group of Glu116 and the center of mass of the thiol group of GSH.

ensembles. However, the variant AgGSTE5 seems unfit to bind DDT as shown by the molecular docking calculations, and should be unable to catalyze DDT. Therefore, the present simulations suggest that among the three variants, only AgGSTE2-I114T/F120L is catalytically efficient at DDT conversion. It implies that only AgGSTE2-I114T/F120L is expected to confer DDT resistance to *A. gambiae* consistent with mutational studies of this enzyme in populations of DDT-resistant and susceptible mosquito.[1]

Conclusions

Proteins are phenotypic traits that are retained or lost in populations via evolutionary pressures. It has been previously shown that mutations present in the coding region of the GSTE2 gene from *A. gambiae* have been associated to DDT resistance, and these polymorphisms were used to track DDT resistance.[16] Furthermore, GSTE2 gene displays the highest level of conservation among all the GSTEs in *Anopheles* species whereas the GSTE5 gene is evolving under positive selection.[62] The negative selection curbing AgGSTE2 sequence divergence favors its DDT dehydrochlorinase activity, the highest ever reported for any GST enzyme[13,14] On the other hand, several sites in AgGSTE5 are under positive selection: 4 residues in the N-terminal domain, where the binding of glutathione occurs (G-site), and 14 residues in the substrate-binding site (H-site).[63] The increased rate of non-similar amino acid replacements in the G-site and H-site strongly suggest that AgGSTE5 is undergoing selective pressure to express dehydrochlorinase activity.[16] The molecular foundations of GST-based insecticide resistance are not completely understood. From a molecular perspective, new functions emerge via mutations that displace the conformational equilibrium toward alternative structures enabling new functions.[64-66] Hence, an increase in mutation rates favoring a given enzymatic activity will result in increased conformational heterogeneity, allowing the same enzyme to accommodate different substrates.[67,68]

We have explored the conformational dynamics of three AgGSTE variants (AgGSTE2, AgGSTE2-I114T/F120L and AgGSTE5) in an attempt to identify the molecular basis of their distinct DDTase activity. The present simulations show that in presence of GSH, the three variants undergo a sequence of conformational rearrangements involving helices H4, H5 and H9 which lead either to the occlusion of the G-site by helix H2, or the opening of the G-site due to the conformational rearrangement of the C-terminal region into an antiparallel β-sheet with the C-terminus of helix H4. The formation of such β-sheet precludes the helix H4 from bending over the G-site, leaving GSH less constrained conformationally. During the occlusion of the G-site, Glu116 (Glu120 in GSTE5) approaches the thiol group of GSH, which shifts away from Ser12, the putative catalytic base. In the AgGSTE2-I114T/F120L and AgGSTE5 ensembles, the distance between the carboxylate group of Glu116 and the thiol group of GSH is within a distance compatible with an elimination reaction previously postulated for the conversion of DDT into DDE.[17] Because Glu116 is expected to have a much lower pKa than Ser12, it could be a more likely candidate to act as a base in the activation of GSH. The present simulations indicate that AgGSTE2-I114T/F120L confer DDT resistance to *A. gambiae* on account of its higher efficiency in the activation of GSH compared to AgGSTE2, and higher binding affinity for DDT compared to AgGST5. In addition, we have built and validated a structural model of the GSTE5 isoform suitable for applications such as prediction of the effect of protein mutations and structure-guided virtual screening of drugs that could help to minimize insecticide resistance in the main vector of malaria in Africa.

Supplementary Information

The AgGSTE5 model structure was deposited in the Protein Model Data Base (https://bioinformatics.cineca.it/) under PMDB ID PM0077788. Supplementary data (sequence alignment, Ramachandran plots, secondary structure time evolution and topologies) are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

References

1. Mitchell, S. N.; Rigden, D. J.; Dowd, A. J.; Lu, F.; Wilding, C. S.; Weetman, D.; Dadzie, S.; Jenkins, A. M.; Regna, K.; Boko, P.; Djogbenou, L.; Muskavitch, M. A.; Ranson, H.; Paine, M. J.; Mayans, O.; Donnelly, M. J.; *PLoS One* **2014**, *9*, e92662.

2. Hayes, J. D.; Pulford, D. J.; *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 445.

3. Oakley, A.; *Drug Metab. Rev.* **2011**, *43*, 138.

4. Frova, C.; *Biomol. Eng.* **2006**, *23*, 149.

5. Sheehan, D.; Meade, G.; Foley, V. M.; Dowd, C. A.; *Biochem. J.* **2001**, *360*, 1.

6. Enayati, A. A.; Ranson, H.; Hemingway, J.; *Insect Mol. Biol.* **2005**, *14*, 3.

7. Kostaropoulos, I.; Papadopoulos, A. I.; Metaxakis, A.; Boukouvala, E.; Papadopoulou-Mourkidou, E.; *Insect. Biochem. Mol. Biol.* **2001**, *31*, 313.

8. Armstrong, R.; *Chem. Res. Toxicol.* **1997**, *10*, 2.

9. French-Constant, R. H.; Daborn, P. J.; le Goff, G.; *Trends Genet.* **2004**, *20*, 163.

10. Hemingway, J.; Field, L.; Vontas, J.; *Science* **2002**, *298*, 96.

11. Corbel, V.; N'Guessan, R.; Brengues, C.; Chandre, F.; Djogbenou, L.; Martin, T.; Akogbeto, M.; Hougard, J. M.; Rowland, M.; *Acta Trop.* **2007**, *101*, 207.

12. David, J. P.; Strode, C.; Vontas, J.; Nikou, D.; Vaughan, A.; Pignatelli, P. M.; Louis, C.; Hemingway, J.; Ranson, H.; *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4080.

13. Djouaka, R.; Irving, H.; Tukur, Z.; Wondji, C. S.; *PLoS One* **2011**, *6*, e27760.

14. Ranson, H.; Rossiter, L.; Ortelli, F.; Jensen, B.; Wang, X.; Roth, C. W.; Collins, F. H.; Hemingway, J.; *Biochem. J.* **2001**, *359*, 295.

15. Ortelli, F.; Rossiter, L. C.; Vontas, J.; Ranson, H.; Hemingway, J.; *Biochem. J.* **2003**, *373*, 957.

16. Ayres, C. F.; Muller, P.; Dyer, N.; Wilding, C. S.; Rigden, D. J.; Donnelly, M. J.; *PLoS One* **2011**, *6*, e29237.

17. Wang, Y. J.; Qiu, L.; Ranson, H.; Lumjuan, N.; Hemingway, J.; Setzer, W. N.; Meehan, E. J.; Chen, L. Q.; *J. Struct. Biol.* **2008**, *164*, 228.

18. Winayanuwattikun, P.; Ketterman, A. J.; *Biochem. J.* **2004**, *382*, 751.

19. Winayanuwattikun, P.; Ketterman, A. J.; *J. Biol. Chem.* **2005**, *280*, 31776.

20. Adcock, S. A.; McCammon, J. A.; *Chem. Rev.* **2006**, *106*, 1589.

21. van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glattli, A.; Hunenberger, P. H.; Kastenholz, M. A.; Ostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 4064.

22. van Gunsteren, W. F.; Bakowies, D.; Burgi, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Glattli, A.; Hansson, T.; Oostenbrink, C.; Peter, C.; Pitera, J.; Schuler, L.; Soares, T. A.; Yu, H. B.; *Chimia* **2001**, *55*, 856.

23. Chothia, C.; Lesk, A. M.; *EMBO J.* **1986**, *5*, 823.

24. Dunbrack, R. L.; *Curr. Opin. Struct. Biol.* **2006**, *16*, 374.

25. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J. H.; Zhang, Z.; Miller, W.; Lipman, D. J.; *Nucleic Acids Res.* **1997**, *25*, 3389.

26. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C.; *Nucleic Acids Res.* **2003**, *31*, 3381.

27. Henikoff, S.; Henikoff, J. G.; *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10915.

28. Brenner, S.; *Nature* **1987**, *329*, 496.

29. Chung, S. Y.; Subbiah, S.; *Structure* **1996**, *4*, 1123.

30. Sali, A.; Blundell, T. L.; *J. Mol. Biol.* **1993**, *234*, 779.

31. Baker, D.; Sali, A.; *Science* **2001**, *294*, 93.

32. Moxon, C.; Wills, B.; *Adv. Exp. Med. Biol.* **2008**, *609*, 131.

33. Rohl, C. A.; Strauss, C. E. M.; Chivian, D.; Baker, D.; *Proteins: Struct., Funct., Genet.* **2004**, *55*, 656.

34. Fernandez-Fuentes, N.; Oliva, B.; Fiser, A.; *Nucleic Acids Res.* **2006**, *34*, 2085.

35. Fiser, A.; Do, R. K. G.; Sali, A.; *Protein Sci.* **2000**, *9*, 1753.

36. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T.; *Bioinformatics* **2006**, *22*, 195.

37. Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T.; *Nat. Protoc.* **2009**, *4*, 1.

38. Melo, F.; Feytmans, E.; *J. Mol. Biol.* **1998**, *277*, 1141.

39. Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C.; *Proteins: Struct., Funct., Genet.* **2000**, *40*, 389.

40. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J.; *J. Comput. Chem.* **1998**, *19*, 1639.

41. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S.; *J. Comput. Chem.* **2007**, *28*, 1145.

42. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J.; *J. Comput. Chem.* **2009**, *30*, 2785.

43. Sanner, M. F.; *J. Mol. Graphics Modell.* **1999**, *17*, 57.

44. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A.; *J. Comput. Chem.* **1986**, *7*, 230.

45. Goodford, P. J.; *J. Med. Chem.* **1985**, *28*, 849.

46. Soares, T.; Goodsell, D.; Ferreira, R.; Olson, A. J.; Briggs, J. M.; *J. Mol. Recognit.* **2000**, *13*, 146.

47. Soares, T. A.; Goodsell, D. S.; Briggs, J. M.; Ferreira, R.; Olson, A. J.; *Biopolymers* **1999**, *50*, 319.

48. Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F.; van Gunsteren, W. F.; *Eur. Biophys. J.* **2005**, *34*, 273.

49. Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F.; *J. Comput. Chem.* **2004**, *25*, 1656.

50. Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F.; *Eur. Biophys. J.* **2011**, *40*, 843.

51. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E.; *J. Chem. Theory Comput.* **2008**, *4*, 435.

52. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C.; *J. Comput. Chem.* **2005**, *26*, 1701.

53. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., ed.; Reidel: Dordrecht, 1981, pp. 331.

54. van Gunsteren, W. F.; Berendsen, H. J. C.; *Mol. Simul.* **1988**, *1*, 173.

55. Nosé, S.; Klein, M. L.; *Mol. Phys.* **1983**, *50*, 1055.

56. Hoover, W. G.; *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695.

57. Parrinello, M.; Rahman, A.; *J. Appl. Phys.* **1981**, *52*, 7182.

58. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M.; *J. Comput. Chem.* **1997**, *18*, 1463.

59. Miyamoto, S.; Kollman, P. A.; *J. Comput. Chem.* **1992**, *13*, 952.

60. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G.; *J. Chem. Phys.* **1995**, *103*, 8577.

61. Darden, T.; York, D.; Pedersen, L.; *J. Phys. Chem.* **1993**, *98*, 10089.

62. Humphrey, W.; Dalke, A.; Schulten, K.; *J. Mol. Graphics* **1996**, *14*, 33.

63. Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C.; *Proteins: Struct., Funct., Bioinf.* **1993**, *17*, 412.

64. Allardyce, C. S.; McDonagh, P. D.; Lian, L. Y.; Wolf, C. R.; Roberts, G. C.; *Biochem. J.* **1999**, *343*, 525.

65. Caccuri, A. M.; Antonini, G.; Board, P. G.; Parker, M. W.; Nicotra, M.; lo Bello, M.; Federici, G.; Ricci, G.; *Biochem. J.* **1999**, *344*, 419.

66. Caccuri, A. M.; Antonini, G.; Nicotra, M.; Battistoni, A.; lo Bello, M.; Board, P. G.; Parker, M. W.; Ricci, G.; *J. Biol. Chem.* **1997**, *272*, 29681.

67. Liu, S.; Zhang, P.; Ji, X.; Johnson, W. W.; Gilliland, G. L.; Armstrong, R. N.; *J. Biol. Chem.* **1992**, *267*, 4296.

68. Widersten, M.; Bjornestedt, R.; Mannervik, B.; *Biochemistry* **1996**, *35*, 7731.

69. Ketterman, A. J.; Saisawang, C.; Wongsantichon, J.; *Drug Metab. Rev.* **2011**, *43*, 253.