Current Research in Structural Biology 4 (2022) 246–255

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Biochemical and structural analysis of a cytosolic sulfotransferase of the malaria vector *Anopheles gambiae* overexpressed in the reproductive tissues

Arianna Esposito Verza [a,1], Riccardo Miggiano [a], Fabrizio Lombardo [b], Carmine Fiorillo [b], Bruno Arcà [b], Beatrice Purghé [a,1], Erika Del Grosso [a], Ubaldina Galli [a], Menico Rizzi [a], Franca Rossi [a,*]

[a] University of Piemonte Orientale, DSF Department of Pharmaceutical Sciences, Largo Donegani, 2, Novara, Italy
[b] Sapienza University of Rome, Department of Public Health and Infectious Diseases – Division of Parasitology, Piazzale Aldo Moro, 5, Rome, Italy

ARTICLE INFO

Handling Editor: Glaucius Oliva

Keywords:
Cytosolic sulfotransferase
SULT
Mosquito
Crystal structure
Differential mRNA expression
Reproductive system

ABSTRACT

The temporary or permanent chemical modification of biomolecules is a crucial aspect in the physiology of all living species. However, while some modules are well characterised also in insects, others did not receive the same attention. This holds true for sulfo-conjugation that is catalysed by cytosolic sulfotransferases (SULT), a central component of the metabolism of endogenous low molecular weight molecules and xenobiotics. In particular, limited information is available about the functional roles of the mosquito predicted enzymes annotated as SULTs in genomic databases. The herein described research is the first example of a biochemical and structural study of a SULT of a mosquito species, in general, and of the malaria vector *Anopheles gambiae* in particular. We confirmed that the AGAP001425 transcript displays a peculiar expression pattern that is suggestive of a possible involvement in modulating the mosquito reproductive tissues physiology, a fact that could raise attention on the enzyme as a potential target for insect-containment strategies. The crystal structures of the enzyme in alternative ligand-bound states revealed elements distinguishing *Ag*SULT-001425 from other characterized SULTs, including a peculiar conformational plasticity of a discrete region that shields the catalytic cleft and that could play a main role in the dynamics of the reaction and in the substrate selectivity of the enzyme. Along with further *in vitro* biochemical studies, our structural investigations could provide a framework for the discovery of small-molecule inhibitors to assess the effect of interfering with *Ag*SULT-001425-mediated catalysis at the organismal level.

1. Introduction

Sulfo-conjugation is an enzyme-catalysed modification that takes part to the chemical metabolism of a wide array of endogenous compounds, including macromolecules and low molecular weight metabolites (Strott, 2002; Gamage et al., 2006), thus directly or indirectly playing many roles in the biology of most living species. Sulfation is also required for the detoxification or, less frequently, for the bio-activation of numerous xenobiotics, and represents a crucial aspect of the tight connection between the organism and the environment (Glatt, 2000). The reaction is carried out by members of the sulfotransferases protein superfamily, that catalyse the stoichiometric transfer of the $–SO_3$ group from the 3′-phosphoadenosine 5′-phosphosulfate (PAPS) sulfate-donor, to a hydroxyl- or a primary amine-group of the acceptor substrate. The reaction releases the inactive co-substrate (*i.e.* PAP) and the sulfate- or sulfamate-conjugated form of the acceptor molecule (Leyh et al., 2013). Sulfotransferases have been assigned to two broad groups: i) the Golgi-resident enzymes, which typically act on large biomolecules (*e.g.* peptides, lipids and glycosaminoglycans) and ii) the cytosolic sulfotransferases (hereon collectively referred to as “SULTs”), which are responsible for transforming numerous small molecules, including hormones, neurotransmitters and drugs (Allali-Hassani et al., 2007; Günal et al., 2019). By increasing water solubility and excretion from the body, SULTs play a major role in the inactivation and detoxification of endogenous compounds and xenobiotics, thereby representing an important component of phase II metabolism. In a more limited number of cases, on the contrary, the reaction

* Corresponding author. University of Piemonte Orientale DSF - Department of Pharmaceutical Sciences, Largo Donegani, 2 - 28100, Novara, Italy.
E-mail address: franca.rossi@uniupo.it (F. Rossi).
[1] Present affiliation: Esposito Verza A., Max Planck Institute of Molecular Physiology, Dortmund (Germany); Purghé B., University of Piemonte Orientale, Department of Translational Medicine, Novara (Italy).

https://doi.org/10.1016/j.crstbi.2022.07.001
Received 10 March 2022; Received in revised form 30 May 2022; Accepted 11 July 2022

leads to a change of the parent compound that results in the potentiation of the molecule activity or in the generation of conjugates displaying mutagenic and/or carcinogenic potential. Alternatively, sulfo-conjugation can regulate the biosynthesis of a given compound *per se*, and the efficient addressing of the resulting product to target sites, in a temporary inactive form, as it has been shown for several steroids (Glatt, 2000). The reaction can even subvert the physiological activity of a given molecule, as in the case of xanthurenic acid (XA), an end-product of the kynurenine pathway of tryptophan degradation that participates in the control of the central nervous system function (Sathyasaikumar et al., 2017), and that, upon transformation in XA 8-*O*-sulfate by specific SULTs, behaves as natriuretic factor in mammals (Cain et al., 2007; Senggunprai et al., 2008). Considering the multifarious roles of SULTs in humans -and in experimental models of mammalian physiology-, it is not surprising that their characterisation is quite advanced. This also allowed for their clustering on the basis of acceptor substrates specificity, primary sequence conservation, structural similarity, and catalytic activity profiles (as admirably reviewed in (Tibbs et al., 2015)). Conversely, although many insect species produce an indisputable impact on global economy and health, the functional involvement of SULT-catalysed modification in insect biology is less documented. Seminal works proposed the existence of sulfation pathways in insect species that transmit diseases to humans (Yang and Wilkinson, 1973; Slade and Wilkinson, 1974), including the African malaria mosquito *Anopheles gambiae*, which is one of the most efficient vectors of the *Plasmodium falciparum* parasite in sub-Saharan Africa. Although somehow fragmentary, these studies provided support to the role of sulfo-conjugation in the homeostasis and/or disposal of central modulators of the insect endocrine function. More recently, genomics and transcriptomics analyses on *Anopheles* (Bonizzoni et al., 2015) and *Aedes* (Faucon et al., 2015; Marcombe et al., 2013; Cattel et al., 2019) mosquitoes suggested a possible contribution of specific SULTs to mosquito resistance to insecticides. Nonetheless, systematic studies on the expression and biochemical properties of mosquito SULTs are missing and, to the best of our knowledge, only a few insect SULTs have been subjected to extensive *in vitro* studies. This is the case, among Diptera species, for *Dm*ST1, *Dm*ST3 and *Dm*ST4 of the fruit fly *Drosophila melanogaster* (Hattori et al., 2008; Fahmy and Baumgartner, 2013) and, among Lepidoptera, for SULT101A1 of the fall armyworm *Spodoptera frugiperda* (hereon indicated as *Sf*DHR, since it is also endowed with a unique retinol dehydratase activity) (Grün et al., 1996; Vakiani et al., 1998) and for *Bm*ST1 and *Bm*SULT of the silkworm *Bombyx mori* (Hattori et al., 2007; Yamamoto and Liu, 2015). Notably, the preferred acceptor-substrate of *Bm*ST1 is the above mentioned XA (Kushida et al., 2011), a compound that exerts crucial roles in insect species (Han et al., 2007; Rossi et al., 2005, 2006; Lima et al., 2012) and that is known to trigger *Plasmodium* gametogenesis in the midgut of *Anopheles* females, soon after their feeding on parasite-infected blood (Qian et al., 2020). Structural information on insect SULTs is even more scarce, being restricted to the crystal structures of *Sf*DHR in different substrate-bound forms (Pakhomova et al., 2001, 2005).

Several putative SULT family members are encoded in the *A. gambiae* genome and in the present work we focused on the product of the VectorBase AGAP001425 gene. In particular, a targeted analysis of AGAP001425 expression profile during development and in selected adult tissues, confirms previous observations obtained by high throughput transcriptomics studies, strengthening the idea of a possible contribution of the corresponding enzyme to the physiology of the mosquito reproductive organs. We, therefore, performed a biochemical characterisation of a recombinant form of the protein (*Ag*SULT-001425), revealing that it is indeed a genuine SULT that is capable of catalysing the sulfation of standard phenolic substrates. The analysis of the crystal structure of *Ag*SULT-001425, in its substrate-free form and in two different ligand-bound states, points to the conformational plasticity of discrete regions of the protein as the main modulators of the enzyme substrate selectivity and catalysis.

2. RESULTS and DISCUSSION

The AGAP001425 expression profile suggests an involvement of the enzyme in the mosquito reproductive system physiology - Preliminary analyses allowed us to identifying 11 potential members of the SULT family in the *A. gambiae* genome (Supplementary Material, Section 1; Supplementary Figure SF1), including AGAP001425, a gene that, according to existing literature, exhibits a quite peculiar expression pattern, mainly restricted to reproductive organs in both *A. gambiae* adult females and males (Supplementary Material, Section 2; Supplementary Table 1) (Baker et al., 2011; Papa et al., 2017). By adopting a quantitative Real Time PCR (RTqPCR)-based approach, we assessed AGAP001425 transcript abundance in key tissues involved in metabolic modifications (midgut and malpighian tubules) and reproductive organs (ovaries and testis), which were dissected from both male and female mosquitoes. Pre-adult stages, namely 1st-2nd and 3rd-4th instar larvae (L1-L2 and L3-L4 in Fig. 1) and pupae, as well as whole adults of both sexes were also collected. Our analysis revealed that the expression of AGAP001425 was significantly higher in male testes than in any other tissue or stage analysed ($p < 0.05$) (Fig. 1). These results are in good agreement both with microarray data obtained by Baker and colleagues (Baker et al., 2011), who also found the highest expression levels in *A. gambiae* ovaries and testes (Supplementary Table 1), and with the RNA-seq analysis of male reproductive organ in different species of the *A. gambiae* complex, where AGAP001425 transcripts appeared typically more abundant in testes rather than in male accessory glands (Izquierdo et al., 2019). As far as pre-adult stages are concerned, the AGAP001425 mRNA was enriched in the early larval stages (L1-L2 in Fig. 1), while a drop in late larvae (L3-L4)

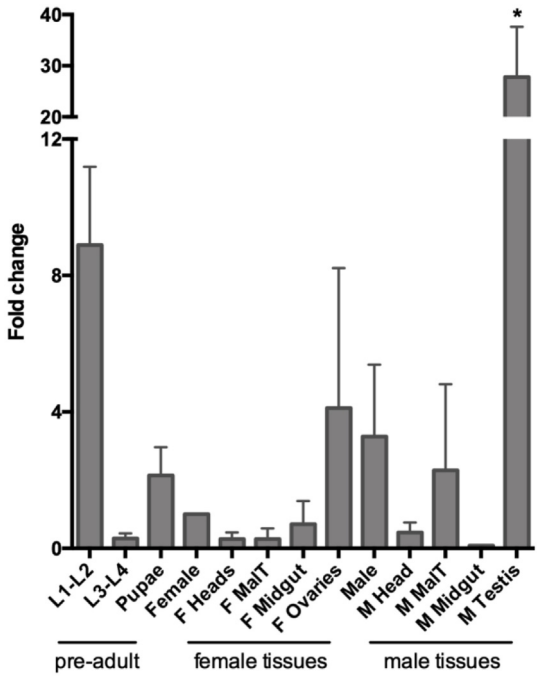
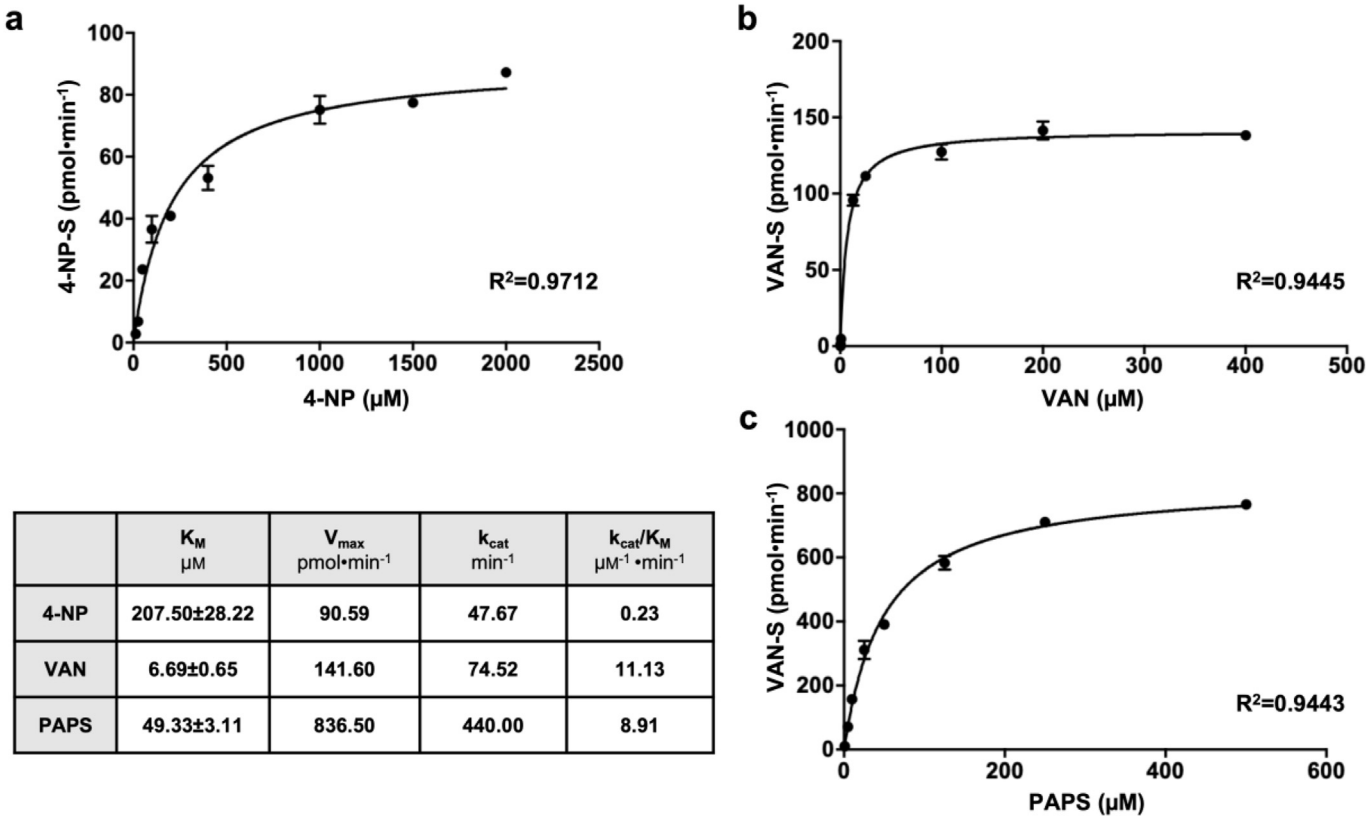


Fig. 1. – AGAP001425 expression analysis. Transcript abundance was evaluated by RTqPCR in pre-adult stages and in dissected tissues from adult male and female mosquitoes, as indicated below the x-axis. Relative standard curve method was applied for quantification, using rpS7 transcript as endogenous reference. Fold change of transcript abundance is calculated setting the average value of whole-body female sample ("Female") as calibrator. Two independent biological replicates were performed and analysed: the average (and the standard deviation) of relative fold change values from both replicates is reported in the graph, and One-way ANOVA followed by "Holm-Sidak's multiple comparisons test" was applied to log10 transformed values to statistically analyse transcript fold changes (*: $p < 0.05$).

was observed. This expression profile is different from the one previously observed by Papa and colleagues (Papa et al., 2017), who found an increase rather than decrease in transcript abundance during the late larval stages; finally, we also detected the AGAP001425 transcript during the pupal stage. Interestingly, the expression of AGAP001425 was recently found modulated during spermatogenesis (Taxiarchi et al., 2019); indeed, transcript abundance appeared to significantly increase from early dividing germline stem cells to meiotic primary spermatocytes, reached a peak in secondary spermatocytes and, finally decreased in post-meiotic spermatids. This expression pattern is outlined in a cluster that includes meiotic genes such as Spo11 (AGAP010898) and could be associated to similar clusters associated with microtubule-based movement, which is essential for sperm motility (Taxiarchi et al., 2019). Overall, our RT-qPCR data are in line with previous microarray- and transcriptomics-based evidence and point to the mosquito reproductive system as a preferential target of this *A. gambiae* SULT. Our hypothesis is consistent, and even strengthened, by the observation that AGAP001425 mRNA levels increase by 4.2-fold in the atrium of *A. gambiae* virgin female mosquitoes upon thoracic injection of 20-hydroxyecdysone, a fundamental hormone in insect physiology, that recently also surfaced as a key regulator of post-mating responses in *A. gambiae* (Gabrieli et al., 2014). The contribution of specific SULTs to the development and functioning of the reproductive system is quite well documented in mammals. In particular, the targeted disruption of the gene encoding the oestrogen sulfotransferase SULT1E1 has been shown to induce structural lesions in adult mice testis, accompanied by reduced sperm production and spermatozoa motility (Qian et al., 2001).

AgSULT-001425 is a bona fide SULT - In light of the evidences reported above, we then shifted our research focus to the predicted AGAP001425-encoded protein. The primary sequence identity to *Dm*ST4 (52%), *Sf*DHR (34%), and *Bm*ST1 (33%), and their clustering together in our phylogenetic analysis (Supplementary Figure SF1), suggest that the mosquito enzyme could have biochemical properties in common with these other insect SULTs. To verify this hypothesis, we over-expressed, in *E. coli*, a variant of the enzyme bearing a 10xHis-tag at its N-terminus, that allows the purification of the recombinant protein (*Ag*SULT-001425) in a single immobilized metal affinity chromatography step; moreover, a size exclusion chromatography-based (SEC) analysis suggested that, in our conditions, *Ag*SULT-001425 is monomeric in solution (Supplementary Figure SF2), similarly to *Sf*DHR but differently from the majority of mammalian SULTs, which assemble into functional dimers. Upon the optimization of *in vitro* assays and analytical procedures to study the enzyme activity (Supplementary Material, section 3; Supplementary Figure SF3), we showed that *Ag*SULT-001425 efficiently catalyses the sulfation of 4-nitrophenol (4-NP) and vanillin (VAN), using PAPS as the co-substrate (Fig. 2). The K_M value of the transformation of 4-NP (K_M^{4-NP} = 207.50 μM) falls within the average range of the ones reported for other insect SULTs (240 μM, 1.3 mM, and 37 μM for *Dm*ST1, *Dm*ST3 and *Dm*ST4 respectively), while the mosquito enzyme shows a higher K_M value for the sulfate-donor PAPS (K_M^{PAPS} = 49.33 μM) compared to the *Drosophila* isoforms (2.1 μM, 24 μM and 1.2 μM for *Dm*ST1, *Dm*ST3 and *Dm*ST4, respectively) (Hattori et al., 2008) and to *Sf*DHR (0.95 μM) (Vakiani et al., 1998). Moreover, *Ag*SULT-001425 displays a higher affinity for VAN over 4-NP (K_M^{VAN} = 6.69 μM), differently from *Dm*ST4, the insect SULT isoform that, based on primary sequence alignments, should represent the closest functionally equivalent enzyme of a non-mosquito dipteran species. The catalytic efficiency (k_{cat}/K_M) of the enzyme towards the tested substrates (Fig. 2) confirms that *Ag*SULT-001425 shows a ~50-fold higher preference for VAN compared to 4-NP, highlighting the 3-methoxy group of vanillin as a potential molecular determinant of the high affinity of the enzyme towards this compound.

Notably, preliminary experiments indicated that *Ag*SULT-001425 is



	K_M μM	V_{max} pmol•min^{-1}	k_{cat} min^{-1}	k_{cat}/K_M μM^{-1} •min^{-1}
4-NP	207.50±28.22	90.59	47.67	0.23
VAN	6.69±0.65	141.60	74.52	11.13
PAPS	49.33±3.11	836.50	440.00	8.91

Fig. 2. – *Ag*SULT-001425 activity characterisation. Michaelis-Menten kinetics of the reaction catalysed by *Ag*SULT-001425 in the presence of **a**, 4-NP (12.5, 25, 50, 100, 200, 400, 1000, 1500, 2000 μM), **b**, VAN (0.0125, 0.125, 0.625, 12.5, 25, 50, 100, 200, 400 μM) and **c**) PAPS (1, 5, 10, 25, 50, 100, 125, 250, 500 μM). For each point, the analytical standard deviation is reported.

not active towards XA, which is among the preferred acceptor-substrates of the lepidopteran *Bm*ST1 that, in turn, is inactive towards VAN and 4-NP (Kushida et al., 2011). A systematic analysis aimed at the identification of the potential endogenous and exogenous sulfate-acceptor substrates of *Ag*SULT-001425 has been undertaken.

The overall structure of AgSULT-001425 - We extensively explored *Ag*SULT-001425 crystallisation conditions to obtain crystals suitable for soaking trials in the presence of different potential ligands, followed by X-ray diffraction experiments, that allowed us to solve the high-resolution crystal structure of the enzyme in three states: in the absence of substrates (*Ag*SULT, PDB ID:7R0O), in complex with the sulfate-acceptor vanillin only (*Ag*SULT:VAN, PDB ID:7R0S), and in complex with both vanillin and the inactive co-substrate PAP (*Ag*SULT:VAN/PAP, PDB ID:7R0U) (Table 1 and Supplementary Figure SF4). Similar to what has been reported for other SULTs, the mosquito enzyme adopts a roughly globular fold (Fig. 3 a). The small domain consists of a 3-stranded antiparallel β-sheet, four α-helices and connecting loops. Interestingly, when compared to other SULTs, the very N-terminal region of *Ag*SULT-001425 (region 2–26) contain additional secondary structure elements that have been observed only in the crystal structure of *Sf*DHR (Supplementary Figure SF5) and that have been correctly predicted in the currently available *in silico* models of the *Dm*ST1, *Dm*ST2 and *Dm*ST4 (Supplementary Figure SF6) (Jumper et al., 2021). The large domain is organized around a 4-stranded parallel β-sheet sandwiched between α-helices, which is highly conserved among SULTs, with the notable exception of the enzymes of *Schistosoma* species (for a recent structural analysis of *Schistosoma* SULTs see (Taylor et al., 2017)). The juxtaposition of the two domains creates an elongated cleft that hosts the contiguous acceptor substrate- and PAPS-binding pockets, with the catalytic centre in the middle, and that represents a hallmark of the molecular architecture of SULTs. Such a deep, elongated crevice is shielded from the bulk solvent by the region encompassing residues 274–312. A sharp bending within this active site cap ideally splits it in a first segment, closely facing the large domain (in orange in Fig. 3), followed by a mostly random coiled moiety (in green) that resembles an adjustable molecular seat belt for securing the substrates at their corresponding binding sites, as it will be further discussed. *Ag*SULT-001425 and *Sf*DHR have another structural element in common: an "U"-shaped α-helical lid that is instead absent - or much more limited in size - in all other SULTs whose experimental structures have been solved thus far. The lid protrudes out the external surface of the large domain to approach the distal end of the acceptor substrate-binding pocket, where it also establishes contacts to the N-side of the molecular seat belt (Fig. 3 b and Supplementary Figure SF7). This observation corroborates the hypothesis of a participation of the lid to regulate the entry and/or fixing of the acceptor substrate molecule into the corresponding pocket of the enzyme (Pakhomova et al., 2005), by possibly modulating the phenomenon more strictly than the structurally equivalent, shorter and flexible regulatory loop of prototypical mammalian SULTs could do. Primary sequence alignments (Supplementary Figure SF8) and, when available, the comparative analysis of *in silico* predicted 3D models (Supplementary Figure SF6), reveals that the few biochemically characterized SULTs of different insect species could possess a shorter or less structured lid, with the exception of *Dm*ST4, thus, again validating the phylogenetic co-clustering of AGAP001425 with *Dm*ST4 (Supplementary Figure SF1).

It is important to point out that the lid has been proposed to have a main role in the exquisitely unique PAPS-dependent retinol dehydratase activity carried out by *Sf*DHR, a reaction that proceeds through the formation of a short-living retinyl-sulfate intermediate (Vakiani et al., 1998; Pakhomova et al., 2001, 2005). However, at present, the relative contribution of the amino acid composition, the size and the structural flexibility of the lid to the two reactions that co-exist in *Sf*DHR is not fully understood. For this reason, an analogous retinol dehydratase activity of *Ag*SULT-001425 cannot be postulated on the simple basis of our structural comparison analysis, and still awaits a direct *in vitro* demonstration.

The structures of AgSULT-001425 in alternative states: a contribution to the description of the molecular choreography accompanying the catalysis - The analysis of optimally superposed *Ag*SULT and *Ag*SULT:VAN structures shows that the binding of a small sulfate-acceptor molecule at the corresponding active site pocket, is not accompanied by any drastic repositioning of the side chains of residues projecting into the catalytic cleft, nor of the protein backbone (Fig. 4). This is probably due to the fact that vanillin is a simple phenolic compound that could easily permeate the acceptor substrate-binding site, without requiring any substantial conformational rearrangement of the protein. Interestingly, in the *Ag*SULT structure, the highly-conserved lysine and histidine residues that play a key role in catalysis (*i.e.* Lys150 and His152 of the mosquito enzyme), participate to the coordination of a water molecule that is also at hydrogen bond distance to the glycerol molecule that we modelled inside the acceptor-substrate binding pocket (Fig. 4 a; Supplementary figure SF4). Consequently, these two crucial residues appear to be already properly oriented to assist the transfer of the $-SO_3$ group to the incoming acceptor substrate, even in the absence of PAPS. Our hypothesis finds support in the analysis of the

Table 1
Data collection and refinement statistics.

Dataset	*Ag*SULT	*Ag*SULT::VAN	*Ag*SULT::VAN/PAP
Data collection			
Wavelength (Å)	0.9677	0.9763	0.9159
Resolution range (Å)	42.74–2.0 (2.072–2.0)	51.74–2.1 (2.175–2.1)	59.55–2.5 (2.59–2.5)
Unit cell constants			
a (Å)	78.773	78.933	78.24
b (Å)	92.147	92.135	91.82
c (Å)	101.761	102.462	103.45
α, β, γ (°)	90	90	90
Space group	P212121	P212121	P212121
N° of molecules in asu	2	2	2
N° of reflections	99487 (9858)	289828 (26028)	191611 (17191)
N° of unique reflections	50299 (4977)	44293 (4361)	26123 (2462)
R_{pim} (%)	5.2 (39.4)	4.86 (16.5)	3.07 (14.92)
R_{merge} (%)	5.2 (39.4)	11.47 (36.65)	7.74 (37.12)
R_{meas} (%)	7.3 (55.7)	12.48 (40.26)	8.35 (40.1)
I/σ	8.4 (1.95)	9.11 (4.79)	14.21 (4.45)
Completeness (%)	99.03 (99.4)	99.81 (99.61)	98.69 (94.90)
Multiplicity	2.0 (2.0)	6.5 (6.0)	7.3 (7.0)
Refinement			
R-factor (%)	19.22 (31.55)	18.54 (23.31)	21.16 (26.50)
R-free (%)	22.60 (37.75)	21.81 (26.53)	24.68 (33.51)
N° of water molecules	306	177	84
N° of non-water ligands			
Glycerol	6	2	3
inorganic ion	2 chlorides	2 chlorides	–
Vanillin	–	2	2
PAP	–	–	2
Average B-factor (Å^2)	28.05	28.85	48.30
macromolecules	27.80	28.77	48.34
Ligands	53.35	44.67	54.06
Solvent	29.65	28.14	39.53
Average RMSD			
bond (Å)	0.014	0.014	0.015
angle (°)	1.78	1.87	1.90
Ramachandran statistics			
Residues (%)			
in most favoured regions	98.79	98.20	97.46
in additional allowed regions	1.21	1.80	2.54
in disallowed regions	0.00	0.00	0.00
Rotamer outliers (%)	0.82	1.14	2.77
Clashscore	2.52	2.16	4.29

Values in parentheses refer to the highest-resolution shell.
[a] asu, asymmetric unit.
[b] R-free (%), calculated from 5% of the data (Brünger, 1992).
[c] RMSD, Root Mean Square Deviation (from ideal).

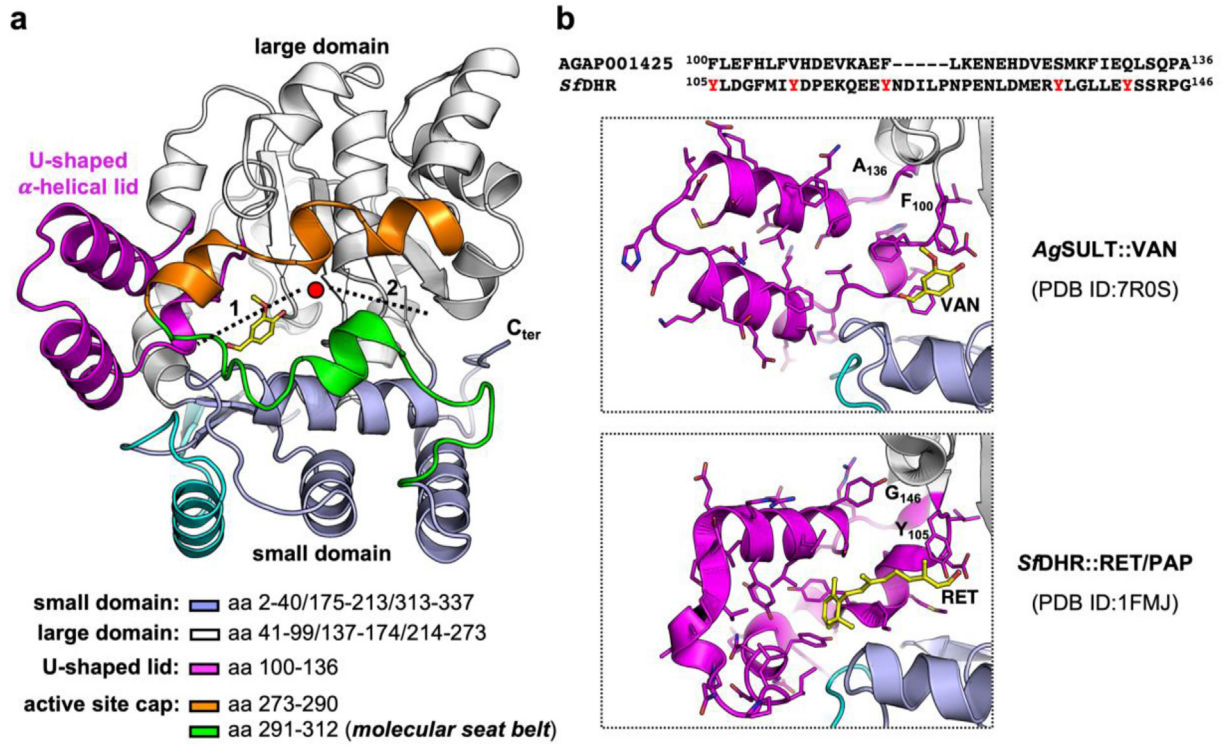


Fig. 3. - Overall structure of *Ag*SULT-001425 - a. Representation of the molecular architecture of *Ag*SULT-001425, as observed in the crystal structure of the enzyme in complex with vanillin (VAN). The catalytic cleft is illustrated by a dotted line, with the red dot labelling the catalytic centre that stands between the adjacent acceptor substrate- and PAPS-binding pockets (1 and 2); the N-terminal region encompassing residues 2–26 appears in cyan. **b**. Primary sequence alignment of the region of the lid of *Ag*SULT-001425 (UniProt: Q7PXJ0) and *Sf*DHR (UniProt: Q26490); the tyrosine residues characterising the *S. frugiperda* enzyme are shown in red; the cartoons illustrate the conformation adopted by the lid in the indicated structures (only the lid residues are shown as sticks); the region corresponding to the active site cap has been omitted and the VAN and retinol (RET) molecules are rendered as sticks. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

*Ag*SULT:VAN structure (Fig. 4 b) that reveals that the vanillin 4-hydroxyl group (*i.e.* the one which can undergo sulfation) is indeed at hydrogen bond distance to both His152 and Lys150, without requiring the repositioning of their side chain. In turn, Lys150 engages the vanillin 3-methoxy group, overall keeping the molecule in the observed orientation, that probably explains why vanillin is a much more preferred substrate of the enzyme compared to 4-NP, which lacks a bioisostere substituting group on its phenolic moiety.

Other residues, lining the inner surface of the acceptor substrate-binding pocket, fix vanillin in place, including Thr192, that establishes a hydrogen bond to the substrate aldehyde group, and a number of aromatic residues that, together with Leu186, create a sort of hydrophobic cage fitting the substrate benzyl ring.

A striking difference among the acceptor-substrate binding pocket of *Ag*SULT-001425 and *Sf*RHD, is represented by the absence, in the mosquito enzyme, of several tyrosine residues that, instead, characterise *Sf*RHD (Fig. 3 b) and that have been proposed to allow proton removal from the retinol cyclohexenyl moiety during the dehydration reaction (Pakhomova et al., 2001, 2005). At present, we cannot demonstrate if other residues of *Ag*SULT-001425 could compensate for this function, nor if the residues projecting into the catalytic site could undergo repositioning only upon bulky molecules binding. However, the size, the conformation and the chemical environment of the *Ag*SULT-001425 sulfate-acceptor substrate-binding pocket, together with the local similarity to the equivalent site of other SULTs, suggest that more structurally complex compounds could in principle behave as ligands of the mosquito enzyme.

The crystal structure of *Ag*SULT-001425 in ternary complex with both substrates, shows that the enzyme undergoes conformational rearrangements upon PAP binding, affecting solvent-exposed regions of the protein, mainly at the level of discrete segments of the molecular seat belt. The PAP molecule is held in place through a series of interactions that are strongly conserved in other structurally characterized SULTs (Fig. 4 c). In detail, the nitrogen atom at position 3 and the NH_2 group of the adenine moiety of PAP are at close bond distance to the phenolic group of Tyr237 and the carbonyl oxygen of Leu271, respectively. Interestingly, the side chain of Trp72, which belongs to a conserved nucleotide-binding α-helix, results approximately 90° rotated compared to the conformation displayed by the same residue in the *Ag*SULT and *Ag*SULT:VAN structures, resulting in a perfect π-stacking to the PAP nitrogenous base (Fig. 4 d). Thr70 and Thr71 engage the 5′-phosphate-moiety of PAP through their side chain hydroxyl- and backbone amide groups, in a highly-coordinated network that also includes Arg67. On the opposite side, the PAP 3′-phosphate group interacts with the side chain of conserved residues Arg174, Ser182 and Arg302, with the participation of the Asn303 backbone amide group. It is worth of mentioning that these latter two residues belong to a short α-helix (residues 298–303) that is observed in the middle of the molecular seat belt in the *Ag*SULT and *Ag*SULT:VAN structures (Fig. 5 a), but that is absent in the *Ag*SULT:VAN/PAP one (Fig. 5 b). We could figure out that the binding of PAPS to a not yet catalytically competent binding site induces the unfolding of the α-helix, to promote the repositioning of some residues that allows proper co-substrate fixing and effective catalysis. Such a strictly required induced-fit event translates into the reshaping of the catalytic cleft as a whole. In particular, in the *Ag*SULT:VAN/PAP structure, the repositioning of the Phe300 residue creates a bottleneck that acts like a gate between the two substrate-binding pockets (Fig. 5). Conversely, when present (*i.e.* in the absence of PAPS or PAP), the rigid α-helix could limit an unrestrained flexibility of the molecular seat belt, thus contributing to hold it in the conformation observed in both the *Ag*SULT and

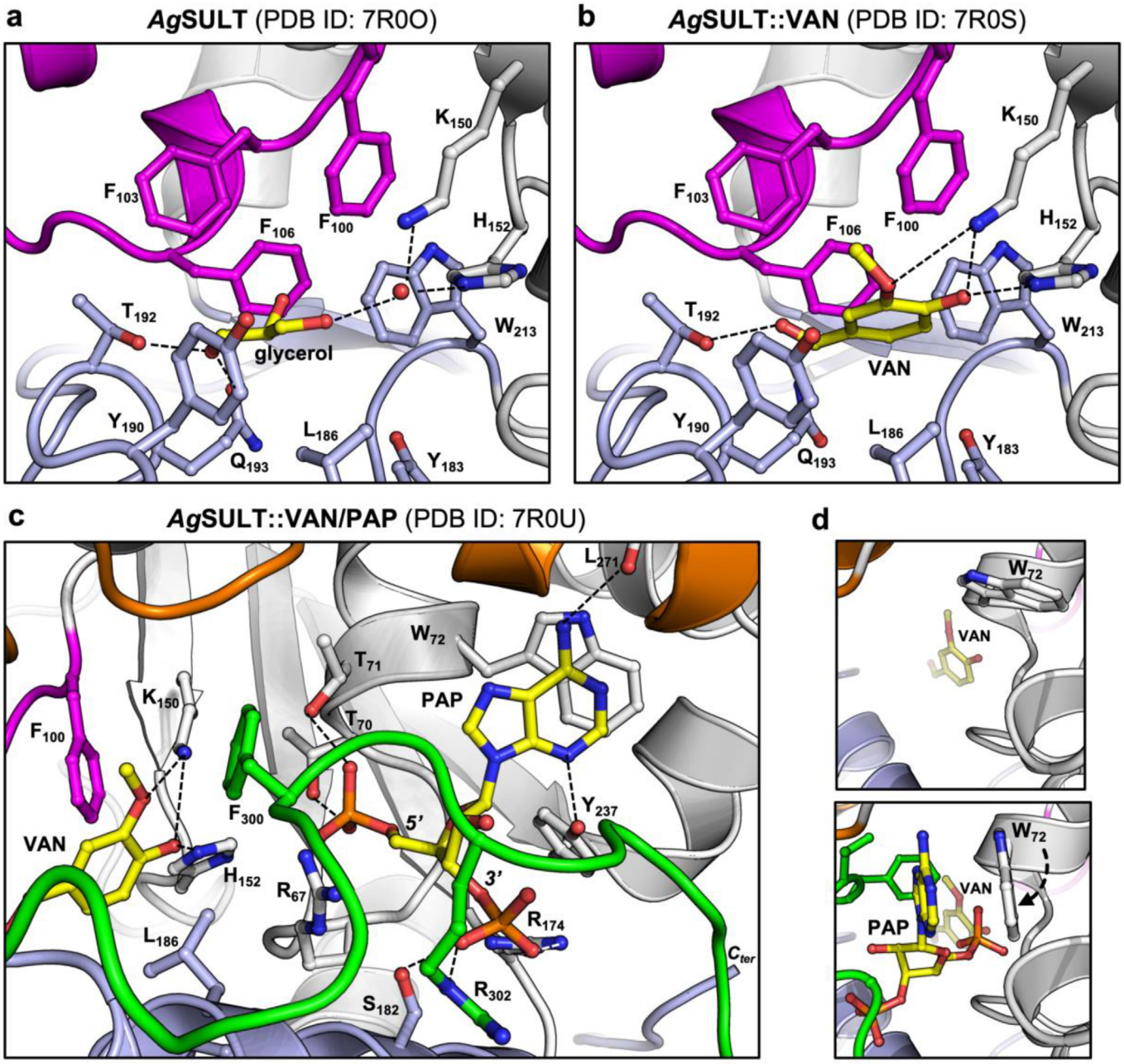


Fig. 4. - *Ag*SULT-001425 substrate-binding sites - Close up of the main molecular contacts that fix one molecule of glycerol (**a**), or vanillin (VAN) (**b**), or the inactive form of the PAPS co-substrate (PAP) (**c**) inside the corresponding binding sites; in **c** part of the acceptor substrate binding pocket is also shown. In **a** and **b,** the active site cap has been omitted. Images shown in d illustrate the flipping of the Tyr72 side chain upon PAP-binding. In all panels, the ligand molecules (yellow) and relevant protein residues are rendered as ball-and-stick, and the protein backbone appears as a cartoon. The colours used to depict the main structural domains and elements are the same than those used in Fig. 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

*Ag*SULT:VAN structures, finally allowing the substrates to more easily access the catalytic cleft. Interestingly, in the seat belt of the available AlphaFold-based models of *Dm*ST1, *Dm*ST2 and *Dm*ST4 (Supplementary Figure SF6), an equivalent short α-helix is absent, probably indicating that these models better depict the co-evolutionarily relevant PAPS/PAP-bound forms of the enzymes than the ligand-free ones. Among published SULT structures, only human dehydroepiandrosterone (DHEA) sulfotransferase displays an equivalent α-helix, although, in that case, the presence of this secondary structure element was proposed to be associated to an inhibition mechanism (Rehse et al., 2002). Overall, we consider the conformational plasticity of the cap as a whole, and of the molecular seat belt in particular, as the main choreographer of *Ag*SULT-001425 dynamics during catalysis.

3. Conclusions

The RTqPCR-based targeted analysis of tissue- and developmental expression of AGAP001425 confirms a significant enrichment of the corresponding transcript in the *Anopheles* reproductive system, especially in males. This observation is of particular interest and, if corroborated by *in vivo* studies, would highlight AGAP001425 as a potentially interesting target for the development of new insect-control strategies. Sequence comparison and phylogenetic analyses revealed that the AGAP001425-encoded protein is the *A. gambiae* SULT most closely related to *Dm*ST4, *Bm*ST1 and *Sf*DHR. However, the variability of residues building up the SULTs acceptor-substrate binding site somehow hampers the *a priori* identification of those responsible for substrate specificity, by merely

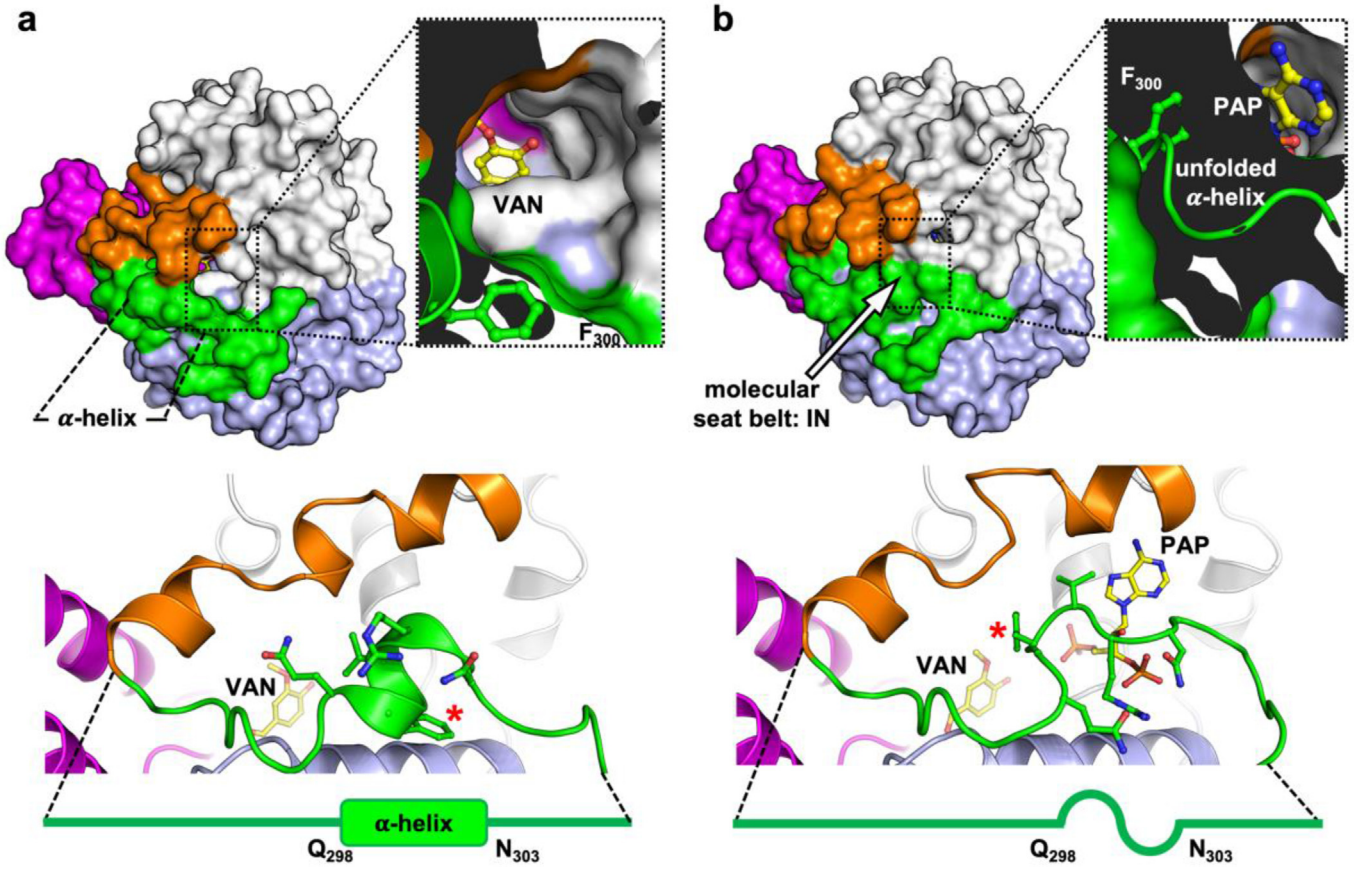


Fig. 5. - The re-shaping of the catalytic cleft upon PAP binding - Upper panels: surface representations of optimally superposed *Ag*SULT:VAN (**a**) and *Ag*SULT:VAN/PAP (**b**) structures, as observed from the co-substrate binding-site entrance; the images framed by a dotted line are cross-sectional views of the region of the protein that is indicated by a dotted frame. **Lower panels:** cartoons depicting the conformation adopted by the molecular seat belt in *Ag*SULT:VAN (**a**) and *Ag*SULT:VAN/PAP (**b**), which highlight the structural plasticity of this protein segment; the red asterisk tags the Phe300 residue of the molecular seat belt α-helix, which undergoes unfolding at PAP binding. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

basing on primary sequence comparison analyses; and, in facts, in many instances, sequence similarity does not strictly correlate with similar substrate selectivity (Tibbs et al., 2015). The reasons behind those discrepancies could be manifold. First, primary sequence identity between any couple of insect SULTs is relatively low (30–35%), and values up to 50% can be recorded only by comparing enzymes of insects belonging to the same order (*e.g.* Diptera, such as *D. melanogaster* and *A. gambiae*, or Lepidoptera, as in the case of *S. frugiperda* and *B. mori*) (Supplementary Figure SF8). Secondly, induced-fit phenomena, documented in almost every study on SULTs, including our one on *Ag*SULT-001425, could occur to a different extent when considering different isozymes. Our analyses suggest that the conformational transitions involving the molecular seat belt could play a primary role in tuning the series of events that controls the trafficking of substrates and reaction products in to and out from the active sites during catalysis. Obviously, we cannot exclude that the binding of bulkier ligands (either alternative physiological sulfate-acceptor substrates or enzyme inhibitors) could promote much more extensive structural rearrangements, or could simultaneously affect the repositioning or the remodelling of additional structural elements, the U-shaped helical lid among them. However, taken as a whole, our results build up an experimental framework to start a methodical investigation of SULT-mediated modification of endogenous low molecular weight molecules and xenobiotics in *Anopheles*, a largely uncovered aspect of the mosquito ecology.

4. Materials and methods

4.1. *Reagents* - *all chemicals were from Sigma-Aldrich, unless otherwise specified*

Mosquito Rearing and tissues collection. *Anopheles coluzzii* mosquitoes (GA-CAM strain, *Anopheles gambiae* "M" form, originally collected in Cameroon) were reared in the insectary of the Department of Public Health and Infectious Diseases at Sapienza University of Rome under standard conditions (28 °C, 80% relative humidity, 12:12 h light:dark photoperiod). Mosquito cycle maintenance was achieved by sugar feeding on cotton swabs with a 10% sucrose solution placed at the top of the cage and by means of membrane feeding using defibrinated mutton blood. Adult female (not yet blood-fed) and male mosquitoes 2–4 days post-emergence (dpe), were used for tissue dissections. Two independent biological replicates (mosquitoes from different generations) were performed. The following samples were collected for each replicate: L1-L2 instar larvae (N = 125); L3-L4 instar larvae (N = 40); early and late pupae (N = 30); heads (N = 20), midguts (N = 25), malpighian tubules (N = 25), ovaries (N = 25) and whole bodies (N = 10) from adult females; heads (N = 20), midguts (N = 25), malpighian tubules (N = 25), testis (N = 25) and whole bodies (N = 10) from adult males. Samples were collected in PBS (Phosphate-Buffered Saline) and stored at −80 °C until use.

RNA extraction and Real Time PCR. Total RNA was extracted from

samples using Trizol Reagent (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's protocols and finally resuspended in DNAse-RNAse-free ultrapure H_2O (MilliQ purified). RNA quantity and quality were assessed by both spectrophotometry (Take3 module of the multi-well plate reader BioTek SynergyHT and GEN5™ software) and agarose gel electrophoresis. For each sample, 1, 2 or 4 μg of total RNA were treated with TURBO™ DNase (Ambion TURBO DNA-free™ kit), following manufacturer's instruction, to remove genomic DNA contamination. The treatment efficacy was verified through standard PCR, without Reverse Transcription, designed to amplify the *An. coluzzii* rpS7 transcript (AGAP010592), which encodes the 40S ribosomal protein S7. Aliquots of DNase-treated total RNA (200, 400 or 800 ng) were used for first-strand cDNA synthesis using the SuperScript™ II Reverse Transcriptase (Invitrogen) and Oligo$(dT)_{18\text{-}20}$ primers. The cDNA samples were diluted to 10 ng/μL and were used as a template for quantitative real-time PCR. The relative quantification method was applied and included a standard curve for the target AGAP001425 transcript, and for the transcripts encoded by the endogenous reference gene, *rpS7*, in the amplification reactions. The cDNA samples used for the preparation of the standard curves were obtained from entire naïve female mosquitoes following the procedure described before. Serial 1:5 cDNA dilutions were prepared to obtain 5 points for each standard curve at the following concentrations: [100 ng/μL], [20 ng/μL], [4 ng/μL], [0.8 ng/μL], and [0.16 ng/μL]. Each qPCR reaction consisted of the following mixture (20 μL total volume): 2 μL of cDNA [10 ng/μL], Cf = [1 ng/μL], 10 μL of 2X PowerUpTM SYBR® Green Master Mix, and 4 μL of forward and reverse primers, Ci = [1 μM], Cf = [200 nM]. The PCR thermal cycle included an initial holding stage of 10 min at 95 °C followed by 40 amplification cycles (15 s at 95 °C, 1 min at 60 °C). A melt curve stage analysis was introduced to assess primer pairs efficiency and to confirm the production of the specific amplicon for each target gene. Each of the two independent biological replicates was analysed in technical duplicate on the same amplification reaction. The relative quantification analysis was conducted by interpolating the Ct value of each candidate with the standard curve and then by normalizing it with the relative value of the rpS7 endogenous reference. The female whole-body sample was set as calibrator for the calculation of fold-change values. The primers used for the PCR amplifications are:

ACO_SULT_F: ATCAAGGAAGGTTGGGCTGA.
ACO_SULT_R: GTCTTGCGTATTGTGTCCGG.
RPS7_F: GTGCGCGAGTTGGAGAAGA.
RPS7_R: ATCGGTTTGGGCAGAATGC.

Graph of Fig. 1 and statistical analysis were performed using GraphPad Prism software. Statistical analysis: one-way ANOVA followed by "Holm-Sidak's multiple comparisons test" was applied to log10 transformed (normalized) values to statistically analyse transcript fold changes (*: $p < 0.05$).

Sub-cloning, protein expression and purification - The DNA corresponding to the AGAP001425 open reading frame was amplified by PCR, using the AGAP001425-PA_pMA-T synthetic construct (Invitrogen) as the template. The PCR product was inserted into the expression plasmid pET-16b (Novagen), by standard procedures, and the protein-encoding region in the resulting pET-*Ag*SULT-001425 construct was verified by sequencing (Eurofins, MWG Operon). Several colonies of *E. coli* BL21(DE3) bacteria, freshly transformed with pET-*Ag*SULT-001425, were dispersed in of 2XYT selective medium to reach an optical density at 600 nm ($O.D._{600}$) of 0.15, and grown at 37 °C till an $O.D._{600}$ = 0.8, under vigorous shaking. The expression of the recombinant protein was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the bacterial culture, followed by a further 5 h incubation, at 30 °C. Induced bacteria were collected by centrifugation, dissolved in buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 1 mM DTT) and disrupted using an ultrasonic processor (Vibra-Cell), on ice. Upon the addition of an EDTA-free commercial protease inhibitor cocktail, the bacterial lysate was centrifuged 30 min at 16000×*g* at 4 °C. The obtained supernatant was added of 10 mM imidazole and incubated under slight rotation in the presence of Ni-NTA agarose resin (Promega), pre-equilibrated following manufacturer's instructions. At the end of the incubation, the protein-laden resin was packed into a disposable column, and extensively washed in buffer A containing 40 mM imidazole. The stepwise elution of the recombinant protein was obtained by fluxing in the column fixed volumes of buffer A, at increasing imidazole concentration. *Ag*SULT-001425-containing eluted fractions were pooled, dialyzed against stock buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT) and concentrated using centrifugal ultrafiltration devices (molecular weight cut-off = 30 kDa) (VivaSpin, Vivascience). The procedure reproducibly yields ≈25 mg of 99% pure recombinant *Ag*SULT-001425 *per* litre of induced bacterial culture. Size exclusion chromatography (SEC) was performed using an Akta FPLC instrumentation and a Superdex® 200 Increase 10/300 GL column (Cytiva), pre-calibrated with standard proteins (following manufacturer's instructions), in 50 mM Tris-HCl, pH 8.0 as the mobile phase. During the entire procedure, the recombinant protein was monitored by standard SDS-PAGE analysis. The protein concentration was determined by the Bradford assay, using bovine serum albumin as the standard.

Enzyme activity characterisation - Assay conditions to test *in vitro* the enzyme activity -in terms of optimal pH and temperature, and $MgCl_2$ requirement-, were optimized by adopting a previously reported LC-MS/MS analytical procedure (Canavesi et al., 2019), with modifications (Supplemental Material, section 3). The K_M value of the *Ag*SULT-001425-catalysed reaction towards the sulfate donor has been determined using a fixed concentration of the sulfate acceptor VAN (100 μM), in the presence of different concentrations of PAPS (1–500 μM). Similarly, the K_M values towards the two acceptor substrates 4-NP and VAN, have been determined using a fixed concentration of PAPS (100 μM) in the presence of different concentrations of 4-NP (12.5 μM - 2 mM) or VAN (12.5 nM - 400 μM). In all cases, the incubation mixture consisted of 50 mM sodium phosphate buffer pH 6.5, 1 mM DTT, 1 mM $MgCl_2$, 1 mg/mL BSA and 80 ng of the recombinant enzyme. After 20 min at 37 °C, the reaction was stopped by the addition of ice-cold acetonitrile (methanol in case of VAN), in 1:1 ratio. Samples were centrifuged at 13000×*g*, at 4 °C, for 10 min and diluted 1:40 prior to be subjected to UPLC-HRMS-based analysis. Each assay was performed in duplicate and three analyses were performed for each sample. The synthesis of the vanillin-sulfate standard is described in the Supplementary Material, section 4.

Crystal structure determination – Crystallisation trials were performed by adopting a robot-assisted (Oryx4; Douglas Instruments), sitting-drop-based spare-matrix strategy, by mixing, in equal proportion, a 10 mg/mL protein solution (in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT) and the crystallisation buffers of several screen kits (Qiagen), in a final droplet volume of 1 μL. Crystals grew in three conditions, invariably in the form of small plates clusters, in 4 weeks at 20 °C. Crystallisation conditions optimization, by means of the vapour diffusion method in hanging drop and micro-seeding, led to the growth of thin plate-shaped single crystals, using the protein at 7 mg/mL, in 1:1 ratio with the optimized crystallisation buffer, in a final droplet volume of 3 μL. In particular, the crystals that allowed us to solve the *Ag*SULT and *Ag*SULT:VAN structures, grew in 25% w/v PEG 3350, 0.3 M NaCl and 0.1 M Bis/Tris at pH 5.5 or 6.5, respectively; the crystal used to solve the *Ag*SULT:VAN/PAP structure grew in 0.1 M KSCN and 30% PEG 2000 MME. Soaking experiments were conducted by incubating crystals into the corresponding crystallisation buffer, supplemented with 1 mM PAPS (*Ag*SULT), or 10 mM vanillin (*Ag*SULT:VAN), or 10 mM vanillin and 10 mM PAP (*Ag*SULT:VAN/PAP), at 20 °C for 24 h, prior to be transferred in the cryogenic-protectant solution (*i.e.* the soaking buffer containing 20% V/V glycerol), mounted on cryo-loops and flash frozen under liquid nitrogen. The X-ray diffraction dataset of the enzyme in a substrate-free form (*Ag*SULT, Table 1) was collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France; detector: Dectris Eiger X 4 M); the datasets of the enzyme in complex with vanillin alone, and vanillin + PAP (*Ag*SULT:VAN and *Ag*SULT:VAN/PAP in Table 1, respectively), were

collected at the Diamond Light Source (Harwell Science and Innovation Campus, Oxfordshire, UK; detectors: Dectris Eiger XE 16 M and Pilatus 6M-F, respectively). Data were integrated by XDS (Kabsch, 2010), scaled and truncated using SCALA and TRUNCATE programs of the CCP4 suite (Winn et al., 2011). The *Ag*SULT structure was solved by molecular replacement (MR), using the program PHASER of the PHENIX suite (McCoy et al., 2007; Adams et al., 2010) and one monomer of the crystal structure of *Sf*DHR in complex with retinol and PAP as the search model (PDB ID:1FMJ) (Pakhomova et al., 2001), upon excluding ligands and water molecules, and truncating to alanine all residues differing among *Sf*DHR and *Ag*SULT-001425 sequences, by using the SCULPTOR function of PHENIX (Bunkoczi and Read, 2011). Automated crystallographic refinement with Phenix.refine (Afonine et al., 2012) and Refmac5 (Murshudov et al., 2011) were alternated to manual model building sessions, using the Coot program (Emsley and Cowtan, 2004). Waters were automatically added by "ARP/wARP Solvent" (Perrakis et al., 1997) and inspected by Coot. Interestingly, the inspection of the density map of the resulting *Ag*SULT structure did not reveal any signal interpretable as the unreacted substrate PAPS (used in the crystal soaking step), nor as its inactive form PAP. An analogous procedure allowed us to solve the structure of *Ag*SULT-001425 in complex with vanillin and in complex with both vanillin and PAP, by using one monomer the *Ag*SULT structure as the search model at the MR-based phasing step. The 10xHis N-terminal tag and the two very C-terminal residues of the recombinant protein are not defined in any of our structures, and they have been consequently omitted from the final models, together with residues 284–285 of the *Ag*SULT structure. Focusing at the enzyme ligand binding pockets, the visual inspection of the *Ag*SULT:VAN and *Ag*SULT:VAN/PAP structures revealed electron density univocally attributable to the compounds used in crystal soaking, while we modelled a glycerol molecule, the cryo-protectant used for data collection, in the active site of *Ag*SULT. Data collection and refinement statistics are given in Table 1. All structure illustrations have been generated using PyMol (DeLano, 2009). The atomic coordinates and structure factors of *Ag*SULT, *Ag*SULT:VAN and *Ag*SULT:VAN/PAP have been deposited in the Protein Data Bank under accession codes 7R0O, 7R0S and 7R0U, respectively.

Funding

This research has been supported by Fondazione Cariplo (grant no 2015-2246) and by the University of Piemonte Orientale. A. E. V. was the recipient of a fellowship granted by Fondazione Novara Sviluppo (Novara, IT).

CRediT authorship contribution statement

Arianna Esposito Verza: Investigation, Data curation, Writing – original draft, Writing – review & editing. **Riccardo Miggiano:** Validation, Data curation, Writing – original draft, Writing – review & editing. **Fabrizio Lombardo:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Carmine Fiorillo:** Investigation. **Bruno Arcà:** Conceptualization, Validation, Writing – original draft, Writing – review & editing. **Beatrice Purghé:** Investigation. **Erika Del Grosso:** Methodology, Data curation, Writing – original draft, Writing – review & editing. **Ubaldina Galli:** Methodology, Validation. **Menico Rizzi:** Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing. **Franca Rossi:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities (beamline ID30B) and Diamond Light Source for time on beamlines I03–I04, under Proposal MX-20221. We wish also to thank P. Serini and M. Calzetta for the help with mosquito rearing and for providing the different mosquito developmental stages, M. Stagnoli and M. Nugnes for technical assistance with enzyme purification and analysis, and I. D'Angelo for fruitful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crstbi.2022.07.001.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, N., Echols, I.W., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. D66, 213–221.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. D68, 352–367.

Allali-Hassani, A., Pan, P.W., Dombrovski, L., Najmanovich, R., Tempel, W., Dong, A., Loppnau, P., Martin, F., Thornton, J., Edwards, A.M., Bochkarev, A., Plotnikov, A.N., Vedadi, M., Arrowsmith, C.H., 2007. Structural and chemical profiling of the human cytosolic sulfotransferases. PLoS Biol. 5, e97.

Baker, D.A., Nolan, T., Fischer, B., Pinder, A., Crisanti, A., Russell, S., 2011. A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, Anopheles gambiae. BMC Genom. 12, 296.

Bonizzoni, M., Ochomo, E., Dunn, W.A., Britton, M., Afrane, Y., Zhou, G., Hartsel, J., Lee, M.-C., Xu, J., Githeko, A., Fass, J., Yan, G., 2015. RNA-seq analyses of changes in the Anopheles gambiae transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. Parasites Vectors 8, 474.

Brünger, A.T., 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–474.

Bunkoczi, G., Read, R.J., 2011. Improvement of molecular-replacement models with Sculptor. Acta Crystallogr. D Biol. Crystallogr. D67, 303–312.

Cain, C.D., Schroeder, F.C., Shankel, S.W., Mitchnick, M., Schmertzler, M., Bricker, N.S., 2007. Identification of xanthurenic acid 8-O-beta-D-glucoside and xanthurenic acid 8-O-sulfate as human natriuretic hormones. Proc. Natl. Acad. Sci. U. S. A. 104, 17873–17878.

Canavesi, R., Miggiano, R., Stella, M., Galli, U., Rossi, F., Rizzi, M., Del Grosso, E., 2019. Study of Anopheles gambiae 3-hydroxykynurenine transaminase activity and inhibition by LC-MS/MS method. J. Pharm. Biomed. Anal. 173, 154–161.

Cattel, J., Faucon, F., Le Péron, B., Sherpa, S., Monchal, M., Grillet, L., Gaude, T., Laporte, F., Dusfour, I., Reynaud, S., David, J.P., 2019. Combining genetic crosses and pool targeted DNA-seq for untangling genomic variations associated with resistance to multiple insecticides in the mosquito Aedes aegypti. Evol. Appl. 13, 303–317.

DeLano, W.L., 2009. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, 2002.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Fahmy, K., Baumgartner, S., 2013. Expression analysis of a family of developmentally-regulated cytosolic sulfotransferases (SULTs) in Drosophila. Hereditas 150, 44–48.

Faucon, F., Dusfour, I., Gaude, T., Navratil, V., Boyer, F., Chandre, F., Sirisopa, P., Thanispong, K., Juntarajumnong, W., Poupardin, R., Chareonviriyaphap, T., Girod, R., Corbel, V., Reynaud, S., David, J.P., 2015. Identifying genomic changes associated with insecticide resistance in the dengue mosquito Aedes aegypti by deep targeted sequencing. Genome Res. 25, 1347–1359.

Gabrieli, P., Kakani, E.G., Mitchell, S.N., Mameli, E., Want, E.J., Mariezcurrena Anton, A., Serrao, A., Baldini, F., Catteruccia, F., 2014. Sexual transfer of the steroid hormone 20E induces the postmating switch in Anopheles gambiae. Proc. Natl. Acad. Sci. U. S. A. 111, 16353–16358.

Gamage, N., Barnett, A., Hempel, N., Duggleby, R.G., Windmill, K.F., Martin, J.L., McManus, M.E., 2006. Human sulfotransferases and their role in chemical metabolism. Toxicol. Sci. 90, 5–22.

Glatt, H., 2000. Sulfotransferases in the bioactivation of xenobiotics. Chem. Biol. Interact. 129, 141–170.

Grün, F., Noy, N., Hämmerling, U., Buck, J., 1996. Purification, cloning, and bacterial expression of retinol dehydratase from Spodoptera frugiperda. J. Biol. Chem. 271, 16135–16138.

Günal, S., Hardman, R., Kopriva, S., Mueller, J.W., 2019. Sulfation pathways from red to green. J. Biol. Chem. 294, 12293–12312.

Han, Q., Beerntsen, B.T., Li, J., 2007. The tryptophan oxidation pathway in mosquitoes with emphasis on xanthurenic acid biosynthesis. J. Insect Physiol. 53, 254–263.

Hattori, K., Hirayama, M., Suzuki, H., Hamamoto, H., Sekimizu, K., Tamura, H., 2007. Cloning and expression of a novel sulfotransferase with unique substrate specificity from Bombyx mori. Biosci. Biotechnol. Biochem. 71, 1044–1051.
Hattori, K., Motohashi, N., Kobayashi, I., Tohya, T., Oikawa, M., Tamura, H.O., 2008. Cloning, expression, and characterization of cytosolic sulfotransferase isozymes from Drosophila melanogaster. Biosci. Biotechnol. Biochem. 72, 540–547.
Izquierdo, A., Fahrenberger, M., Persampieri, T., Benedict, M.Q., Giles, T., Catteruccia, F., Emes, R.D., Dottorini, T., 2019. Evolution of gene expression levels in the male reproductive organs of Anopheles mosquitoes. Life Sci. Alliance 2, e201800191.
Jumper, J., Evans, R., Pritzel, A., et al., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589.
Kabsch, W., 2010. XDS. *Acta Crystallogr. D Biol. Crystallogr.* D66, 125–132.
Kushida, A., Horie, R., Hattori, K., Hamamoto, H., Sekimizu, K., Tamura, H., 2011. Xanthurenic acid is an endogenous substrate for the silkworm cytosolic sulfotransferase, bmST1. J. Insect Physiol. 58, 83–88.
Leyh, T.S., Cook, I., Wang, T., 2013. Structure, dynamics and selectivity in the sulfotransferase family. Drug Metab. Rev. 45, 423–430.
Lima, V.L., Dias, F., Nunes, R.D., Pereira, L.O., Santos, T.S., Chiarini, L.B., Ramos, T.D., Silva-Mendes, B.J., Perales, J., Valente, R.H., Oliveira, P.L., 2012. The antioxidant role of xanthurenic acid in the Aedes aegypti midgut during digestion of a blood meal. PLoS One 7, e38349.
Marcombe, S., Paris, M., Paupy, C., Bringuier, C., Yebakima, A., Chandre, F., David, J.P., Corbel, V., Despres, L., 2013. Insecticide-driven patterns of genetic variation in the dengue vector Aedes aegypti in Martinique island. PLoS One 8, e77857.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.
Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. D67, 355–367.
Pakhomova, S., Kobayashi, M., Buck, J., Newcomer, M.E., 2001. A helical lid converts a sulfotransferase to a dehydratase. Nat. Struct. Biol. 8, 447–451.
Pakhomova, S., Buck, J., Newcomer, M.E., 2005. The structures of the unique sulfotransferase retinol dehydratase with product and inhibitors provide insight into enzyme mechanism and inhibition. Protein Sci. 1, 176–182.
Papa, F., Windbichler, N., Waterhouse, R.M., Cagnetti, A., D'Amato, R., Persampieri, T., Lawniczak, M.K.N., Nolan, T., Papathanos, P.A., 2017. Rapid evolution of female-biased genes among four species of Anopheles malaria mosquitoes. Genome Res. 27, 1536–1548.
Perrakis, A., Sixma, T.K., Wilson, K.S., Lamzin, V.S., 1997. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. Acta Crystallogr. D Biol. Crystallogr. 53, 448–455.
Qian, Y.M., Sun, X.J., Tong, M.H., Li, X.P., Richa, J., Song, W.C., 2001. Targeted disruption of the mouse estrogen sulfotransferase gene reveals a role of estrogen metabolism in intracrine and paracrine estrogen regulation. Endocrinology 142, 5342–5350.
Qian, P., Zhang, C., Zhong, C., Su, X.Z., Yuan, J., 2020. An intracellular membrane protein GEP1 regulates xanthurenic acid induced gametogenesis of malaria parasites. Nat. Commun. 11, 1764.
Rehse, P.H., Zhou, M., Lin, S.-H., 2002. Crystal structure of human dehydroepiandrosterone sulphotransferase in complex with substrate. Biochem. J. 364, 165–171.
Rossi, F., Lombardo, F., Paglino, A., Cassani, C., Miglio, G., Arcà, B., Rizzi, M., 2005. Identification and biochemical characterization of the Anopheles gambiae 3-hydroxykynurenine transaminase. FEBS J. 272, 5653–5662.
Rossi, F., Garavaglia, S., Giovenzana, G.B., Arcà, B., Li, J., Rizzi, M., 2006. Crystal structure of the Anopheles gambiae 3-hydroxykynurenine transaminase. Proc. Natl. Acad. Sci. U. S. A. 103, 5711–5716.
Sathyasaikumar, K.V., Tararina, M., Wu, H.Q., Neale, S.A., Weisz, F., Salt, T.E., Schwarcz, R., 2017. Xanthurenic acid formation from 3-hydroxykynurenine in the mammalian brain: neurochemical characterization and physiological effects. Neuroscience 367, 85–97.
Senggunprai, L., Yoshinari, K., Shimada, M., Yamazoe, Y., 2008. Involvement of ST1B subfamily of cytosolic sulfotransferase in kynurenine metabolism to form natriuretic xanthurenic acid sulfate. J. Pharmacol. Exp. Therapeut. 327, 789–798.
Slade, M., Wilkinson, C.F., 1974. Degradation and conjugation of Cecropia juvenile hormone by the southern armyworm (*Prodenia eridania*). Comp. Biochem. Physiol. 49B, 99–103.
Strott, C.A., 2002. Sulfonation and molecular action. Endocr. Rev. 23, 703–732.
Taxiarchi, C., Kranjc, N., Kriezis, A., Kyrou, K., Bernardini, F., Russell, S., Nolan, T., Crisanti, A., Galizi, R., 2019. High-resolution transcriptional profiling of Anopheles gambiae spermatogenesis reveals mechanisms of sex chromosome regulation. Sci. Rep. 9, 14841.
Taylor, A.B., Roberts, K.M., Cao, X., Clark, N.E., Holloway, S.P., Donati, E., Polcaro, C.M., Pica-Mattoccia, L., Tarpley, R.S., McHardy, S.F., Cioli, D., LoVerde, P.T., Fitzpatrick, P.F., Hart, P.J., 2017. Structural and enzymatic insights into species-specific resistance to schistosome parasite drug therapy. J. Biol. Chem. 292, 11154–11164.
Tibbs, Z.E., Rohn-Glowacki, K.J., Crittenden, F., Guidry, A.L., Falany, C.N., 2015. Structural plasticity in the human cytosolic sulfotransferase dimer and its role in substrate selectivity and catalysis. Drug Metabol. Pharmacokinet. 30, 3–20.
Vakiani, E., Luz, J.G., Buck, J., 1998. Substrate specificity and kinetic mechanism of the insect sulfotransferase, retinol dehydratase. J. Biol. Chem. 273, 35381–35387.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. D67, 235–242.
Yamamoto, K., Liu, M., 2015. Identification and characterization of a new type of cytosolic sulfotransferase in the silkworm Bombyx mori. J. Insect Biotechnol. Sericol. 84, 63–68.
Yang, R.S.H., Wilkinson, C.F., 1973. Sulphotransferases and phosphotransferases in insects. Comp. Biochem. Physiol. 46B, 717–726.