structural communications

A micromolar O-sulfated thiohydroximate inhibitor bound to plant myrosinase

Arthur Besle,[a] Xavier Brazzolotto,[a]‡ Arnaud Tatibouët,[b] Deimante Cerniauskaite,[b] Estelle Gallienne,[b] Patrick Rollin[b] and Wim P. Burmeister[a,c]*

[a]UJF–EMBL–CNRS UMI 3265, Unit of Virus Host Cell Interactions, BP 181, France, [b]Institut de Chimie Organique et Analytique, UMR 6005, Associé au CNRS, Université d'Orléans, BP 6759, F-45067 Orléans, France, and [c]Institut Universitaire de France (IUF), 103 Boulevard St Michel, F-75005 Paris, France

‡ Present address: Centre de Recherche du Service de Santé des Armées Emile Pardé, BP 87, 38702 Grenoble, France.

Correspondence e-mail: wpb@embl.fr

The 1.6 Å resolution structure of the micromolar competitive inhibitor S-(N,N-dimethylaminoethyl) phenylacetothiohydroximate-O-sulfate bound to *Sinapis alba* myrosinase, a plant thioglucosidase, is reported. Myrosinase and its substrates, the glucosinolates, are part of the plant's defence system. The sulfate group and the phenyl group of the inhibitor bind to the aglycon-binding site of the enzyme, whereas the N,N-dimethyl group binds to the glucose-binding site and explains the large improvement in binding affinity compared with previous compounds. The structure suggests ways to increase the potency and specificity of the compound by improving the interactions with the hydrophobic pocket of the aglycon-binding site.

1. Introduction

Myrosinase is a unique enzyme which catalyzes the hydrolysis of sulfur-containing secondary metabolites called glucosinolates (**1**), which are usually regarded as 1-thio-β-D-glucosides (Fig. 1). Glucosinolates are simple but fascinating molecules that are characterized by an O-sulfated (Z)-thiohydroximate function, a thio-linked β-D-glucopyrano moiety and a most often hydrophobic aglycon chain, the constitution of which is the sole structural variant in nature. More than 120 different chain structures have been identified in plant species of the *Brassicales* order (Fenwick *et al.*, 1983). In association with myrosinase (EC 3.2.1.147), glucosinolates form a host defence
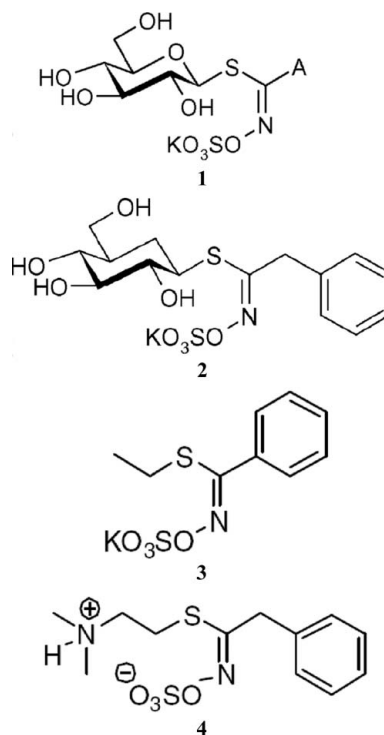
Figure 1
(**1**) General structure of glucosinolates; in glucotropaeolin A is a benzyl group. (**2**) 5a-Carba-glucotropaeolin. (**3**) S-Ethyl phenylacetothiohydroximate-O-sulfate. (**4**) S-(N,N-Dimethylaminoethyl) phenylacetothiohydroximate-O-sulfate (PDB residue name e18).

Table 1
Sample information.

PDB code	2wxd
Crystallization method	2 + 2 µl hanging drops over 500 µl reservoir solution at 293 K
Crystallization solutions	100 mM Tris–HCl pH 8, 68% $(NH_4)_2SO_4$ as precipitant, protein at 3.5 mg ml^{-1} in 20 mM HEPES pH 6.5, 150 mM NaCl, 20 µM $ZnSO_4$
Space group	$C222_1$
Unit-cell parameters (Å, °)	$a = 136.25$, $b = 137.57$, $c = 80.72$, $\alpha = \beta = \gamma = 90$

Table 2
Data-collection and structure-solution statistics.

Values in parentheses are for the outer shell.

Diffraction source	Synchrotron; ESRF beamline ID14eh2
Wavelength (Å)	0.933
Detector	ADSC Quantum4
Temperature (K)	100.0
Resolution range (Å)	68.8–1.60 (1.64–1.60)
No. of unique reflections	99631
Completeness (%)	99.6 (98.4)
Redundancy	4.63 (3.87)
$\langle I/\sigma(I)\rangle$	5.90 (1.71)
R_{merge}	0.08 (0.38)
Data-processing software	*MOSFLM* (Leslie, 1992), *SCALA* (Evans, 1993), *CCP*4 (Collaborative Computational Project, Number 4, 1994)

mechanism in plants. The byproducts that result from myrosinase activity on glucosinolates possess insecticidal, fungicidal and bactericidal properties. *Brassica*-specific insects adapted to this system use various detoxification mechanisms. The crystal structure of *Sinapis alba* myrosinase is known (Burmeister *et al.*, 1997). 2-Fluoroglucotropaeolin has previously been characterized as a mechanism-based inhibitor, leading to a stable 2-fluoroglucosyl intermediate that has been used for structural studies (Burmeister *et al.*, 1997, 2000). Another class of inibitors are nonspecific glucosidase inhibitors with a planar anomeric atom that mimic the transition state (Heightman & Vasella, 1999). These inhibitors bind to the glucose-binding site (Burmeister *et al.*, 2000). Inhibitors targeted against the aglycon-binding site are more recent and to date have shown inhibition in the millimolar range at best (Bourderioux *et al.*, 2005; Aucagne *et al.*, 2000). As the result of optimization of the previous compounds, the inhibitor *S*-(*N,N*-dimethylaminoethyl) phenylaceto-

thiohydroximate-*O*-sulfate (**4** in Fig. 1), with an IC$_{50}$ of 3.3 µM, has recently been synthesized (Cerniauskaite *et al.*, 2009). Here, we report the structure of this inhibitor bound to *S. alba* myrosinase.

2. Methods

Myrosinase was purified from *S. alba* seeds as described previously (Pessina *et al.*, 1990). As it had been stored frozen in 20 mM HEPES pH 6.5 it could not be used directly for crystallization owing to aggregation, hence 150 mM NaCl and 20 µM $ZnSO_4$ were added. The sample was submitted to gel filtration on a Superdex 200 column (GE
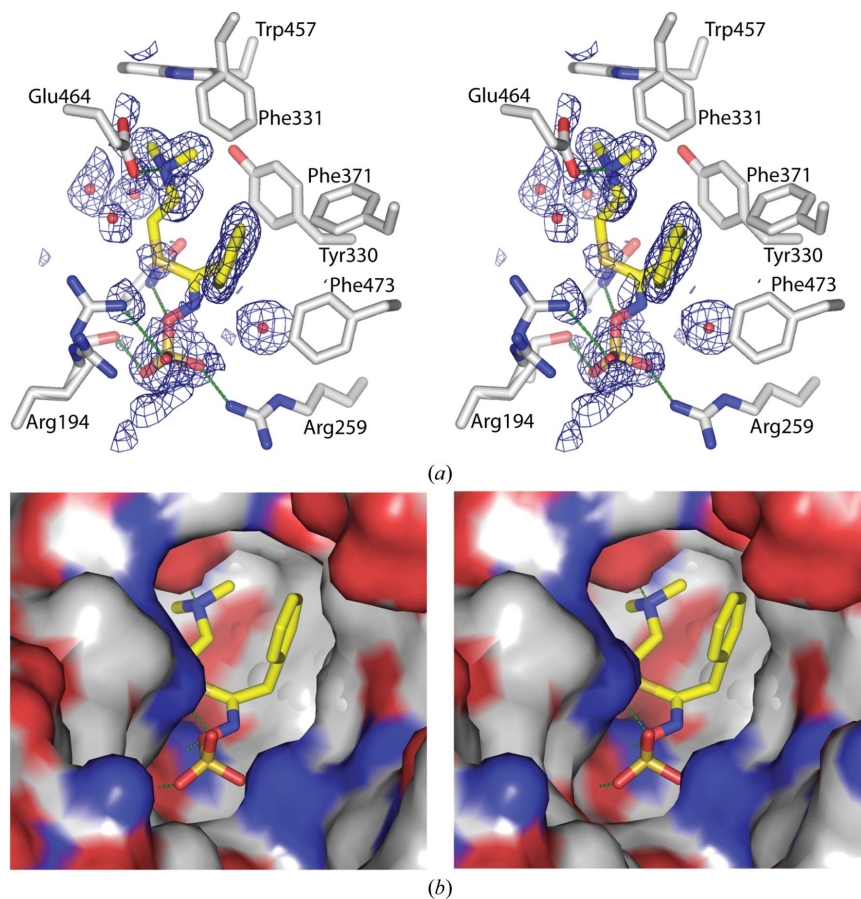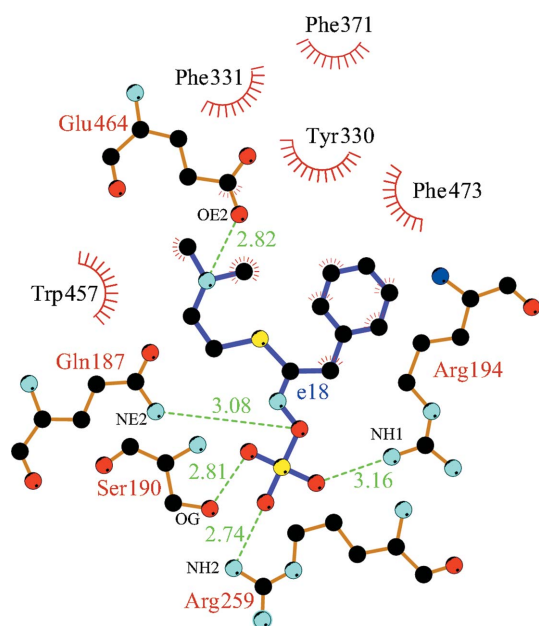

Figure 2
Stereoview of the active site. (*a*) OMIT electron density contoured at 2.5σ and the position of the refined inhibitor (**4** in Fig. 1) together with the interacting residues of the myrosinase structure. Hydrogen bonds are shown as blue dotted lines. Residues Gln187 and Ser190 at the back are not labelled for clarity. Water molecules in the vicinity of the inhibitor are shown as red spheres. (*b*) Surface representation of the active site and the bound inhibitor **4** (yellow C atoms). Hydrogen bonds are shown as blue dotted lines. This figure was prepared with *PyMOL* (http://www.pymol.org).

Healthcare) using the same buffer and the fractions that contained dimeric myrosinase were concentrated. Crystals were obtained under similar conditions (Table 1) as described previously (Burmeister *et al.*, 1997). The O-sulfated thiohydroximate inhibitor (**4**) was obtained as described by Cerniauskaite *et al.* (2009). The inhibitor was solubilized as a 0.5 *M* stock in DMSO and diluted 1:100 with 66% ammonium sulfate in 100 m*M* Tris–HCl pH 8. Part of the inhibitor crystallized in the mother liquor. Crystals were soaked for 30 min before a brief transfer to a cryoprotectant solution containing first 10% and then 20% glycerol, 66% ammonium sulfate, 100 m*M* Tris pH 8. Crystals were frozen in the nitrogen stream and data collection and structure determination are summarized in Tables 2 and 3.

3. Results and discussion

As the crystals were isomorphous to those described previously (Burmeister *et al.*, 1997), structure determination started with rigid-body refinement using *REFMAC* (Murshudov *et al.*, 1997) of the model deposited as PDB entry 1e4m (Burmeister *et al.*, 2000), from which ligands and water molecules in the active site had been removed. The σ_A-weighted $F_o - F_c$ map (Read, 1986) showed additional electron density in the active site. The *N,N*-diaminoethyl group, the phenyl ring and the sulfate moiety of the inhibitor were well defined, as well as side-chain movements. Smaller electron densities corresponded to water molecules in the proximity of the inhibitor (Fig. 2*a*). These electron densities allowed placement of the inhibitor and refinement of the structure with *REFMAC* (Vagin *et al.*, 2004) at 1.6 Å resolution to an R_{cryst} of 14.4% (Table 3). In comparison to the apo structure (PDB code 1e4m), which actually contains a glycerol molecule from the cryoprotectant in the glucose-binding site, structural changes were limited to a few residues of the active site. Glu372 and Glu464 move slightly upon binding of the inhibitor. Arg194 adopts an alternate conformation interacting with the sulfate group of **4**, whereas the other conformation interacts with a sulfate ion from the crystallization solution. This disorder propagates to Asn466, which also adopts two conformations.

Table 3
Structure refinement and model validation.

Values in parentheses are for the outer shell.

Refinement software	*REFMAC* (Murshudov *et al.*, 1997)
Resolution range (Å)	68.8–1.60 (1.64–1.60)
No. of reflections used in refinement	99606
Final overall R factor	0.144
Atomic displacement model	Isotropic
Overall average B factor (Å^2)	12.4
No. of protein atoms	4082
No. of ligand atoms	392
No. of solvent atoms	764
Total No. of atoms	5238
Final R_{work}	0.143 (0.243)
No. of reflections for R_{free}	6886
Final R_{free}	0.164 (0.219)

Lack of solubility and high sulfate content are probable reasons why the inhibitor only has an occupancy of about 80%. The inhibitor only shows weak electron density for the central atoms of the aceto-thiohydroxymate moiety, which is reflected by higher temperature factors for these atoms (32–40 Å^2) than for the peripheral groups (18–25 Å^2); the average temperature factor of the protein atoms in contact with the inhibitor is 15 Å^2. The $F_o - F_c$ map of the final model does not provide any further hints about the origin of the disorder; the remaining difference electron densities correspond to the minor structure in the absence of the inhibitor, with water and glycerol molecules in the same positions as in the apo structure. It is probable that the relative positions of the bound peripheral moieties of the inhibitor are not optimal, leading to conformational flexibility in the central part of the molecule.

The aglycon and sulfate moieties are in very similar positions to those observed previously for 5a-carba-glucotropaeolin (**2**) and *S*-ethyl phenylacetothiohydroximate-*O*-sulfate (**3**) (Bourderioux *et al.*, 2005). The phenyl group binds in the hydrophobic aglycon-binding site formed by Tyr330, Phe331, Phe371 and Phe473. The sulfate moiety binds to Arg194, Arg259, Gln187 and Ser190. The *N,N*-diaminoethyl group forms a salt bridge with the carboxyl group of Glu464; at the same time, the methyl groups contact the hydrophobic residues Trp457, Tyr330 and Phe331 (Figs. 2 and 3). These interactions explain the considerable increase in the IC$_{50}$ of **4** (3.3 μ*M*; Cerniauskaite *et al.*, 2009) compared with the previous compounds **2** (IC$_{50}$ = 0.99 m*M*) and **3** (IC$_{50}$ = 0.58 m*M*; Bourderioux *et al.*, 2005).

From the crystal structures of the previous myrosinase–inhibitor complexes (Bourderioux *et al.*, 2005) it was obvious that 5a-carba-glucotropeaolin (**2**) was not able to bind simultaneously in the glucose- and aglycon-binding sites, leading to a badly placed glucose group (PDB entry 1w9b; Bourderioux *et al.*, 2005). In agreement with this result, the sugar could be omitted, leading to the better binding inhibitor **3**, but the ethyl group replacing the glucose group did not really contribute to binding (PDB entry 1w9d; Bourderioux *et al.*, 2005). The interaction could be considerably improved with the additional *N,N*-dimethylamino group of **4**. The combined positively charged and hydrophobic character makes this tertiary amino group very interesting in inhibitor design. In contrast to the initial rationale in the design of **4** (Cerniauskaite *et al.*, 2009), the *N,N*-dimethylamino group does not mimic an oxonium-ion transition state interacting with the general nucleophile Glu409. Instead, the tertiary amino group in **4** leads to an efficient interaction with the residues Glu464, Trp457 and Phe473 of the glucose-binding site. These hydrophobic residues normally interact with the C5 and C6 atoms of a bound glucose, whereas Glu464 interacts with the hydroxyl groups in the 4- and 6-positions (Burmeister *et al.*, 1997). As the corresponding

glutamic acid residue is widely conserved in β-glucosidases, this approach may be more generally useful in the design of inhibitors directed against this family of enzymes.

Replacement of the phenyl group of the aglycon by other groups in order to create even more potent inhibitors is suggested from the inhibitor–myrosinase complex structure. On the other hand, the non-optimized character of the interaction of **4** suggests that the compound will be nonspecifically active against a broad spectrum of myrosinases. The compound may become a valuable tool in order to study the molecular basis of plant–insect interactions.

References

Aucagne, V., Gueyrard, D., Tatibouët, A., Quinsac, A. & Rollin, P. (2000). *Tetrahedron*, **56**, 2647–2654.

Bourderioux, A., Lefoix, M., Gueyrard, D., Tatibouët, A., Cottaz, S., Arzt, S., Burmeister, W. & Rollin, P. (2005). *Org. Biomol. Chem.* **3**, 1872–1879.

Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). *Structure*, **5**, 663–675.

Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A. & Henrissat, B. (2000). *J. Biol. Chem.* **275**, 39385–39393.

Cerniauskaite, D., Gallienne, E., Karciauskaite, H., Farinha, A., Rousseau, J., Armand, S., Tatibouët, A., Sackus, A. & Rollin, P. (2009). *Tetrahedron Lett.* **50**, 3302–3305.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Evans, P. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.

Fenwick, G. R., Heaney, R. K. & Mullin, W. J. (1983). *Crit. Rev. Food Sci. Nutr.* **18**, 123–201.

Heightman, T. & Vasella, A. (1999). *Angew. Chem. Int. Ed.* **38**, 750–770.

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

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Pessina, A., Thomas, R. M., Palmieri, S. & Luisi, P. L. (1990). *Arch. Biochem. Biophys.* **280**, 383–389.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. & Murshudov, G. N. (2004). *Acta Cryst.* D**60**, 2184–2195.

Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng.* **8**, 127–134.