Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Complexed and ligand-free high-resolution structures of urate oxidase (Uox) from *Aspergillus flavus*: a reassignment of the active-site binding mode

Pascal Retailleau,[a]* Nathalie Colloc'h,[b] Denis Vivarès,[c]‡ Françoise Bonneté,[c] Bertrand Castro,[d] Mohamed El Hajji,[d] Jean-Paul Mornon,[e] Gérald Monard[f] and Thierry Prangé[a,g]

[a]LURE, Centre Universitaire Paris-Sud, Bâtiment 209D, BP 34, 91898 Orsay CEDEX, France, [b]UMR 6551 CNRS, Université de Caen, Centre Cyceron, Boulevard Becquerel, BP 5229, 14074 Caen CEDEX, France, [c]CRMC2 (UPR 7251 CNRS), Campus de Luminy, Case 913, 13288 Marseille CEDEX 09, France, [d]Sanofi–Synthelabo, 82 Avenue Raspail, 94255 Gentilly CEDEX, France, [e]Laboratoire de Minéralogie–Cristallographie de Paris (UMR 7590 CNRS), Universités Paris VI and Paris VII, Campus Boucicaut, 4 Place Jussieu, Case 115, 75252 Paris CEDEX 05, France, [f]Equipe de Chimie et Biochimie Théoriques (UMR 7565 CNRS), Université Henri Poincaré, Nancy I, BP 239, 54506 Vandoeuvre-les-Nancy CEDEX, France, and [g]Laboratoire de Cristallographie et RMN Biologiques (UMR 8015 CNRS), Faculté de Pharmacie, Université Paris V, 4 Avenue de l'Observatoire, 75270 Paris CEDEX 06, France

‡ Present address: Chemical Engineering, 150 Academic Street, University of Delaware, Newark, DE 19716, USA.

Correspondence e-mail: retailleau@lure.u-psud.fr

Received 13 October 2003
Accepted 23 December 2003

PDB References: ligand-free Uox (LF), 1r56, r1r56sf; AZA complex, 1r51, r1r51sf; MUA complex, 1r4s, r1r4ssf; OXC complex, 1r4u, r1r4usf.

High-resolution X-ray structures of the complexes of *Aspergillus flavus* urate oxidase (Uox) with three inhibitors, 8-azaxanthin (AZA), 9-methyl uric acid (MUA) and oxonic acid (OXC), were determined in an orthorhombic space group (*I*222). In addition, the ligand-free enzyme was also crystallized in a monoclinic form ($P2_1$) and its structure determined. Higher accuracy in the three new enzyme–inhibitor complex structures (Uox–AZA, Uox–MUA and Uox–OXC) with respect to the previously determined structure of Uox–AZA (PDB code 1uox) leads to a reversed position of the inhibitor in the active site of the enzyme. The corrected anchoring of the substrate (uric acid) allows an improvement in the understanding of the enzymatic mechanism of urate oxidase.

1. Introduction

Urate oxidase (uricase; EC 1.7.3.3; referred to here as Uox) is an essential enzyme responsible for the first step of the degradation of uric acid into allantoin (Fig. 1). This enzyme occurs widely in nature, with the exception of the higher primates, including man, which lack this enzyme because of several precise mutations and deletions (Wu *et al.*, 1992). This evolution has been attributed to selection pressure, allowing a higher concentration of the powerfully reducing uric acid in order to prevent the occurrence of cancer in long-lived primates (Ames *et al.*, 1981). Urate oxidase from *Aspergillus flavus* has been produced for 25 years by Sanofi and marketed under the trademark Uricozyme. Recently, Sanofi–Synthelabo has obtained worldwide approval for Fasturtec, which contains rasburicase, a recombinant urate oxidase enzyme produced by a genetically modified *Saccharomyces cerevisiae* strain. The therapeutic prescription is for the treatment and prophylaxy of acute hyperuricaemia in order to prevent renal failure in patients with haematological malignancy with a high tumour burden, who are at risk of rapid tumour lysis or shrinkage on initiation of chemotherapy.

Most enzymes that utilize molecular oxygen as a substrate usually require a cofactor to mediate the chemistry with O_2. Analytical data and the crystal structure of Uox (Colloc'h *et al.*, 1997) have demonstrated that the active enzyme contains neither transition metals and organic cofactors nor modified amino acids. The enzyme catalyzes the oxidation of uric acid to 5-hydroxyisourate at physiological or basic pH (Fig. 1). Isotope-labelling studies have shown that O_2 is reduced to H_2O_2 and that the O atom attached to C5 in 5-hydroxyisourate is derived from a molecule of water (Modric *et al.*, 1992). The CO_2 molecule comes from the C6 carbonyl group of the uric acid and is produced during the decarboxylation/degradation in the last step of the reaction (Bentley & Neuberger, 1952).

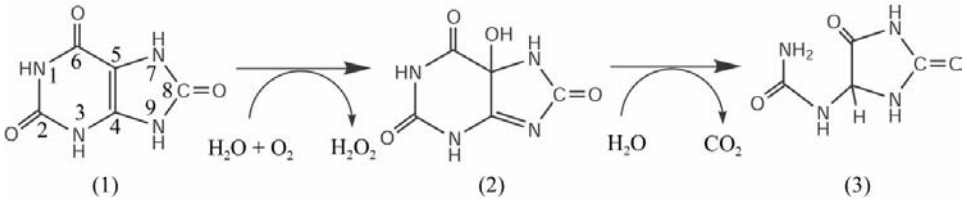

Figure 1
Schematic degradation pathway of uric acid to allantoin through intermediate (2) (5-hydroxyisourate). Only the first step is catalyzed by the enzyme, with H_2O_2 as the by-product.

The mechanism of action of urate oxidase has been extensively explored by Tipton and coworkers (Kahn & Tipton, 1997; Kahn *et al.*, 1997; Imhoff *et al.*, 2003) on the basis of kinetic investigations. All these studies support the intermediacy of 5-hydroxyisourate as the product of the catalyzed step. This intermediate is then degraded to allantoin either chemically or through another enzyme-catalysed step with hydroxyisourate hydrolase, as proposed by Raychaudhuri & Tipton (2002).

The first structure in the urate oxidase family was solved at 2.2 Å resolution in complex with the competitive inhibitor 8-azaxanthin (AZA) using multiple isomorphous replacement and refined to 2.05 Å (Colloc'h *et al.*, 1997). The crystals used to solve this initial structure were grown in six months without precipitating agent. Recently, Bonneté *et al.* (2001) have shown that addition of polyethylene glycol (PEG) to the protein solution accelerated the nucleation, leading to better diffracting crystals (Vivarès & Bonneté, 2002). Data were then recorded from various crystals of Uox in complex with 8-azaxanthin (AZA) and the structure was solved at up to 1.7 Å resolution. This allowed the unambiguous location of the inhibitor in a different orientation from that initially proposed. Furthermore, the structures of complexes with two other inhibitors, 9-methyl uric acid (MUA) and oxonic acid (OXC), in the same space group *I*222 were obtained at equivalent resolutions (1.8 and 1.65 Å, respectively). In parallel, the native structure (ligand-free or LF) was crystallized in a different crystal form ($P2_1$) that diffracted to lower resolution (2.3 Å).

2. Materials and methods

2.1. Crystallizations

Two different protein samples were used. The natural extracted enzyme from *A. flavus* (Uricozyme) and recombinant Uox (rasburicase; trade name Fasturtec in the EU and Elitek in the USA) were prepared and purified at Sanofi–Synthelabo. The complexes with 8-azaxanthin (AZA) and 9-methyl uric acid (MUA) were both crystallized with Uricozyme, while the free enzyme, the oxonic acid (OXC) complex and those pertaining to subsequent studies were obtained from the rasburicase source. All inhibitors were purchased from Sigma–Aldrich. Similar crystallization conditions as described in Bonneté *et al.* (2001) were applied in all cases, using the sitting-drop vapour-diffusion method at room temperature. 5–10 mg ml^{-1} protein in 50 m*M* Tris–HCl pH 8.5 was mixed with the corresponding (0.5–2 mg ml^{-1}) inhibitor solution in the presence of PEG 8000, the concentration of which in the well was within the range 5–8%(*w*/*v*). In the absence of an inhibitor, the crystals display an elongated morphology that is different to the large diamond shapes of the *I*222 crystal forms. They thus appear to be more fragile for mounting into the capillary and diffract to a lower resolution (∼2.3 Å) than the corresponding complexes (≤1.8 Å).

2.2. X-ray data collection and data processing

X-ray data collections were carried out at the DW32 station (Fourme *et al.*, 1992) on the wiggler beamline at the LURE DCI synchrotron, Orsay, France using a wavelength close to 0.97 Å and a 345 mm MAR Research image-plate detector. Temperatures were set in the range 277–283 K using an FTS Systems Inc. air-cooling system. Data were integrated using *DENZO* and scaled independently using *SCALEPACK* from the *HKL* package (Otwinowski & Minor, 1997). Unit-cell parameters were refined independently by the post-refinement technique implemented in the *SCALEPACK* program. The merged intensity data were converted to structure-factor amplitudes using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994; French & Wilson, 1978) and put on an approximate absolute scale using the scale factor taken from a Wilson plot computed within the resolution range of 5 Å d_{max} (see statistics in Table 1).

2.3. Determination of the structures

2.3.1. Orthorhombic *I*222 complexes. The starting model for the refinement of all Uox–inhibitor complexes was taken from the original refined atomic coordinates of the Uricozyme–AZA complex available in the Protein Data Bank (PDB code 1uox). The six C-terminal residues were missing from the model. All solvent and inhibitor molecules were removed from the original model, as well as the acetyl group in the N-terminal position and the free cysteine involved in a disulfide bridge with Cys35. Initial electron-density maps with $|F_{obs}| - |F_{calc}|$ and $2|F_{obs}| - |F_{calc}|$ coefficients were calculated using *SIGMAA* (Read, 1986) and confirmed the unambiguous placement of ligands in the active site, which were incorporated using the graphics program *O* (Jones *et al.*, 1991) and subsequently included in the refinements.

2.3.2. Monoclinic ($P2_1$) ligand-free structure. The ligand-free enzyme crystallizes in space group $P2_1$, with two tetramers per asymmetric unit that are related by a non-crystallographic symmetry twofold axis. Phases were obtained by molecular replacement using the program *EPMR* (Kissinger *et al.*, 1999) and data in the resolution range 15–4 Å. The correlation coefficient after positioning the first tetramer (generated after expansion of the deposited structure

1uox to a full tetramer through the 222 crystallographic symmetry) was 0.466 with an R factor of 0.451; after positioning the second tetramer the correlation coefficient was 0.805 with an R factor of 0.273.

Table 1
Statistics of data collection.

Values in parentheses are for the highest resolution shell.

	Rasburicase, LF (283 K)	Uricozyme–AZA (277 K)	Uricozyme–MUA (283 K)	Rasburicase–OXC (283 K)
Data collection on DW32 (LURE)				
No. of crystals	1	1	2	2
Wavelength (Å)	0.972	0.964	0.968	0.972
Exposure time per frame (s)	120	90	300/240	150/120
Angular increment per frame (°)	1.0	1.0	1.0	1.0
Total rotation range (°)	250	160	150/104	120/120
Crystal-to-detector distance (mm)	345	260	300	250/283
Unit-cell parameter refinement				
Space group	$P2_1$	$I222$	$I222$	$I222$
Unit-cell parameters				
a (Å)	82.59	80.29	80.08	79.86
b (Å)	141.84	96.31	96.46	96.16
c (Å)	134.96	105.56	105.67	105.61
β (°)	92.7			
Data reduction				
No. measured reflections	732235	390325	327273	398514
Resolution limits (Å)	32–2.30	18–1.75	20.0–1.80	20.0–1.65
No. unique reflections	136450	38367	38023	48706
Data completeness (%)	99.2 (98.3)	92.1 (79.6)	100 (99.6)	99.1 (95.9)
Overall $I/\sigma(I)$	12.3 (3.6)	19.8 (4.8)	12.5 (2.5)	13.0 (2.9)
Redundancy	5.2	11.2	9.0	8.2
Overall $R_{\rm sym}$†	0.076 (0.426)	0.065 (0.222)	0.072 (0.245)	0.064 (0.375)
Overall $R_{\rm r.i.m.}$‡	0.084 (0.047)	0.078 (0.273)	0.080 (0.286)	0.068 (0.430)
Overall $R_{\rm p.i.m.}$§	0.035 (0.200)	0.025 (0.139)	0.026 (0.141)	0.022 (0.206)

† $R_{\rm sym}$ is defined as $\sum_{hkl}\sum_i |I_i(hkl) - \overline{I(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\overline{I_i(hkl)}$ is the weighted mean of all observations (after rejection of outliers). ‡ $R_{\rm r.i.m.}$ is the redundancy-independent merging R factor (Weiss & Hilgenfeld, 1997), which is identical to the $R_{\rm meas}$ of Diederichs & Karplus (1997) defined as $\sum_{hkl}(N/N-1)^{1/2}\sum_i |I_i(hkl) - \overline{I_i(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$, with N being the number of times a given reflection has been observed. § $R_{\rm p.i.m.} = \sum_{hkl}(1/N-1)^{1/2}\sum_i |I_i(hkl) - \overline{I_i(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$ is the precision-indicating merging R factor (Weiss & Hilgenfeld, 1997; Weiss, 2001).

Table 2
Refinement statistics.

	Rasburicase, LF	Uricozyme–AZA	Uricozyme–MUA	Rasburicase–OXC
Software	*BUSTER*/*TNT*	*SHELXL*	*BUSTER*/*TNT*	*BUSTER*/*TNT*
No. reflections ($\|F\|$ total)	136244	38362	38023	48706
No. reflections ($\|F\| > 3\sigma$)	121925	34190	30441	38364
No. reflections (used for refinement)	134760	38362	38023	48579
No. reflections (used for $R_{\rm free}$)	13478	3820	3789	4876
No. protein atoms	18904	2374	2386	2378
No. waters	334	201	231	208
Wilson B (Å^2)	47.9	30.0	31.2	26.9
$\langle B\rangle$ for protein atoms (Å^2)	37.2	25.9	26.7	21.7
$\langle B\rangle$ for waters (Å^2)	33.8	41.5	41.6	33.6
$\langle B\rangle$ for inhibitor (Å^2)	—	28.5	24.3	15.3
Resolution range (Å)	10–2.3	10–1.75	20.0–1.80	20.0–1.65
R value† (%)	16.2	17.3	15.4	15.7
$R_{\rm free}$ ‡ (10% $\|F\|$ reserved) (%)	21.5	19.7	18.2	18.2
Weighted r.m.s.d. from ideality				
Bond length (Å)	0.012	0.014	0.011	0.012
Bond angle /dihedral angle (°)	92.3/7.7	92.2/7.8	91.1/8.9	92.9/7.1
Ramachandran plot: % residues in				
Most favoured regions	92.3	92.2	91.1	92.9
Additionally allowed	7.7	7.8	8.9	7.1

† The crystallographic R factor is defined as $\sum |F_{\rm obs}| - |F_{\rm calc}|/\sum |F_{\rm obs}|$ and indicates the accuracy of the model. ‡ The free R factor is a cross-validation residual calculated using 10% of the native data which were randomly chosen and excluded from the refinement.

2.4. Refinement

2.4.1. Orthorhombic *I*222 complexes. Least-squares restrained refinements of atomic coordinates and thermal parameters for the AZA complex were performed using *SHELXL*, part of the *SHELX*97 program package (Sheldrick & Schneider, 1997). Water-molecule positions were updated iteratively using *ARP*/*wARP* (Lamzin & Wilson, 1997; Perrakis *et al.*, 1999). The two other complex structures (with MUA and OXC) were refined using maximum likelihood as implemented in the program *BUSTER* developed by Global Phasing Ltd (Irwin & Bricogne, 1996; Roversi *et al.*, 2000) and interfaced with *TNT* (Tronrud *et al.*, 1987), which handles the stereochemical and geometrical restraints. Parts of the 1uox model that reliably fitted the initial *SIGMAA* electron-density map were kept in order to start the individual parameter refinement. This was performed against all recorded diffracted amplitudes. Truncated pieces and organized parts of solvent were utilized as a model-based non-uniform prior to output low-resolution distributions. Together with the bulk-solvent envelope and the partial model, they provided a statistical description of the scattering from the crystal content at all stages of the refinement process. The *PEAKMAX* and *WATPEAK* programs from *CCP*4 (Collaborative Computational Project, Number 4, 1994) were used to locate the water molecules once the protein model was completed. Waters were inserted in the models if they formed at least one hydrogen bond (between 2.6 and 3.4 Å) with a density peak at least three times the root-mean-square (r.m.s.) level of the $|F_{\rm obs}| - |F_{\rm calc}|$ difference density maps and once the r.m.s. level of the $2|F_{\rm obs}| - |F_{\rm calc}|$ difference density maps. Water molecules with a thermal B factor less than three times the average B factor of the whole structure (protein atoms and solvent) were retained at the end of the refinements.

More than 100 waters were found that were common to the three structures out of a total of more than 200 that were located per monomer.

2.4.2. Monoclinic ($P2_1$) ligand-free structure. Ten cycles of rigid-body refinement were first performed using *BUSTER*/*TNT* on both the position and orientation of the *EPMR* solution (divided into eight copies of domain 1 and eight of domain 2); ten more cycles on their average *B* factors were then performed. The R_{free} factor estimated from 10% of the total amplitudes fell from 29.5 to 26.9%. Residue-by-residue real-space correlation coefficients (RSCC) of the current model with the corresponding *SIGMAA* $2|F_{obs}| - |F_{calc}|$ map formed a normal distribution with average value close to 0.83 and standard deviation of 0.09, accounting for almost 95% of the structure, and a tail representing serious systematic errors from the molecular-replacement model. Removal of the residues (loop 20–25 and regions 220–230 and 250–260 in the eight protomers) in bad agreement with the map (RSCC below 0.7) permitted elimination of the worst phase bias at the start of structure refinement. Twofold strict non-crystallographic symmetry constrained both tetramers to be identical during the early stages of the partial model refinement. Maximum-entropy modulation of the prior prejudice was then applied in order to recover accurate density in the expected regions of missing atoms (Roversi *et al.*, 2000). Note that one C-terminal region could be extended to Ser299 (over 301 residues) for one monomer benefitting from stabilizing crystal-packing

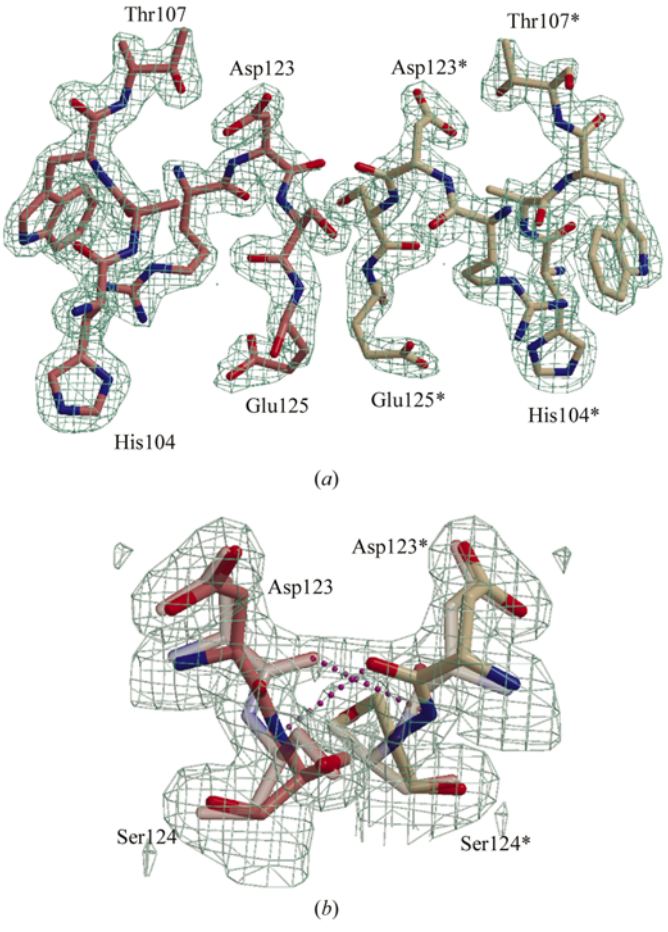

Figure 2
(*a*) Sample of the final electron-density map contoured at 1.5σ. This view was centred on purpose around the region near Asp123 and the twofold-symmetry-related part of the Uox–AZA structure. Apparently, normal tracing of the model leads to a chemical conflict between the peptide carbonyl of Asp123 and its twofold-related counterpart. (*b*) Close-up view at lower contouring (1σ) showing density continuity between the initial positions of the Asp123 carbonyls and new modelling with an alternate flip of the corresponding peptide bonds to alleviate collision.

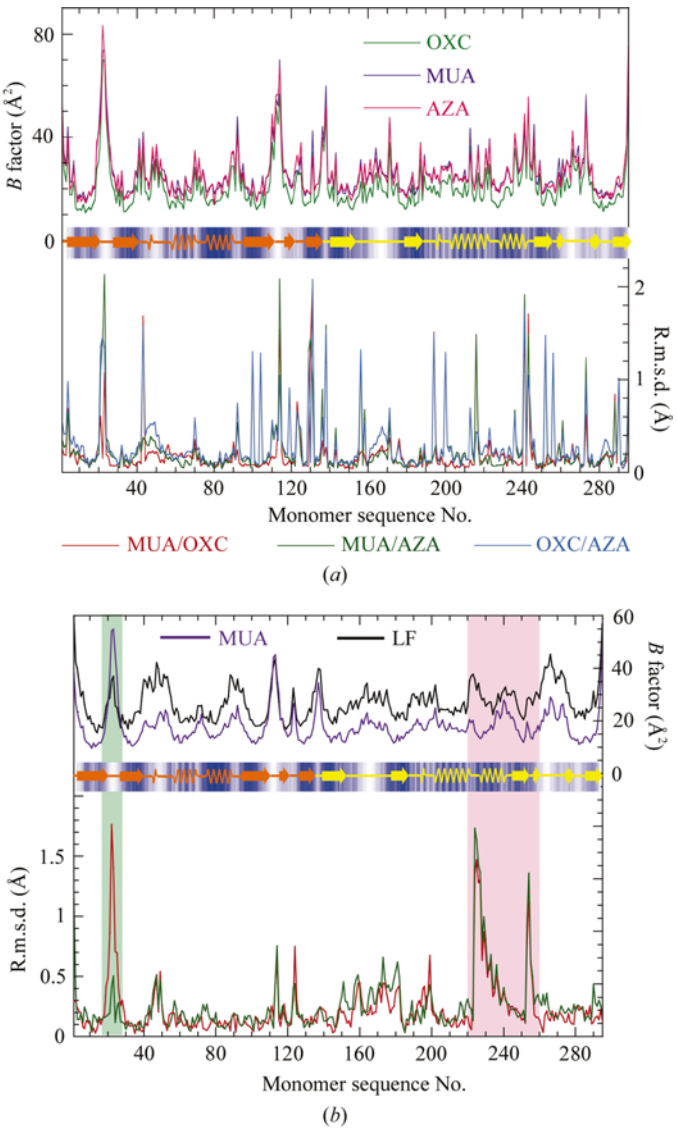

Figure 3
(*a*) Sequence analyses of *B* factors (upper part) and r.m.s. coordinate differences (lower part) *versus* monomeric sequence and secondary structure (central motif coloured as follows: orange, domain *A*; yellow, domain *B*; white background, solvent-accessible residues; blue background, buried residues; as estimated by *PROCHECK*). The two curves represent average values taken for the three analyzed AZA, MUA and OXC complexes of urate oxidase, underscoring their close conformational identities. (*b*) The same graph (upper part, *B* factors; lower part, r.m.s. coordinate differences) for main-chain residues between the *I*222 complex (MUA) and the $P2_1$ ligand-free structure (LF). In the lower part, the two LF protomers (of eight) that present the lowest (green) and the highest (red) r.m.s. coordinate differences with Uox–MUA were selected. The most important differences are highlighted by the two coloured domains and arise from (i) packing constraints (green domain) and (ii) active-site differences (pink domain).

contacts. Once both copies of the tetramer fitted the density with satisfactory resulting R-factor and R_{free} values (21.0 and 23.3%, respectively), waters were then introduced according to the same procedure detailed above. Non-crystallographic symmetry constraints were finally relaxed in order to allow independent descriptions of the eight active-site regions. The final model has a mean real-space correlation of 0.88, with a standard deviation of less than 0.06.

The final refinement statistics for the three complexed monomers and the ligand-free bis-tetramer are summarized in Table 2.

3. Results and discussion

3.1. Crystallographic model refinement and inhibitors bound to protein molecules

3.1.1. AZA, MUA and OXC complexes. Fig. 2 displays a region of the electron-density map representative of the high quality of the electron density throughout the entire unit cell. The three models which could be traced straightforwardly fitting such a density were refined at 1.8 Å resolution or better, with a crystallographic R_{work} value in the range 16–19% and an R_{free} of below 20%. They incorporate 295 amino-acid residues; however, the end of the C-terminal chain could still not be clearly identified in the electron density and thus the final six residues in the sequence were not included in the models. They all contain an acetyl group at the N-terminus and two *cis*-peptide bonds for Pro75 and Pro284, as found originally. As evaluated using *PROCHECK* (Laskowski *et al.*, 1993), all non-glycine residues fall within the energetically favourable regions of the Ramachandran plot (Ramachandran *et al.*, 1963). Superposition of the three complexes reveals very small r.m.s. deviations (0.33 Å between MUA and OXC and 0.41 Å between either MUA or OXC and AZA for all 2300 protein atoms) between the three protomers. Isotropic thermal parameters per residue have comparable values throughout the monomer sequence (Fig. 3*a*). These results confirm that there is basically no structural difference between the extracted and the recombinant protein structures, as expected (Bayol *et al.*, 2002), and that the three inhibitors have similar structural effects. The only noticeable difference between the proteins is located at Cys35. The rasburicase (recombinant enyzme) structure shows no adduct at this particularly solvent-accessible cysteine side chain, while two atoms (an S^{γ} and a C^{β}) could be placed to fit protuberant extra-density suggesting a disulfide bridge between Cys35 and a free cysteine from the culture medium in Uricozyme (the extracted enzyme).

The overall r.m.s. coordinate differences between the new resolution-improved model and the previous AZA-complex structure is 0.41 Å. The main differences arise from correcting side-chain rotamers, particularly when they are long and exposed to the solvent, such as Arg7, Arg105, Arg108, Glu236, Lys114, Lys143, Lys171 and Lys273. The C^{γ} (and/or $C^{\delta1}$) atoms of Leu216, Leu244, Leu252, Leu194 (specifically in OXC) and Pro115 have been repositioned to fit the more accurate electron-density map. The imidazole rings of His104, His200 and His256 have been flipped in order to promote hydrogen-bond formation with the surrounding polar amino-acid side chains. Alternate conformation is no longer observed for His98. In contrast, alternate positions are tentatively proposed for side chains at the surface (Gln131, Gln187, Glu43, Glu246 and Arg241) or in the core (Val130, Ile288, Cys290 and especially in the case of Asp 123; see below).

3.1.2. Ligand-free structure. In this case, two complete tetramers are retained as the asymmetric unit content in space group $P2_1$. The 222 symmetry imposed by the orthorhombic space group no longer holds, but the overall structure remains identical, as illustrated in Fig. 3(*b*), which represents the r.m.s. deviations along the main-chain tracing for two independently refined protomers *versus* the MUA monomer structure.

The overall r.m.s. coordinate difference is 0.46 ± 0.03 Å between the eight non-liganded monomer copies and the new *I*222 MUA-complex structure. Backbone fluctuations that mainly account for coordinate differences are highlighted in Fig. 3(*b*). In addition to the loop 20–25, which is usually disordered in *I*222 structures and becomes stabilized in one particular case (protomer *A*) owing to $P2_1$ crystal-packing constraints (Fig. 3*b*, green area), the linker between H3 and H4 is the region that undergoes the most significant rearrangement as a direct consequence of the absence of inhibitor (Fig. 3*b*, pink zone). However, the backward motion of that chain entails only a marginal expansion of the active-site cavity. Fig. 4 displays a close-up of this region with the new interaction network created in the absence of inhibitor.

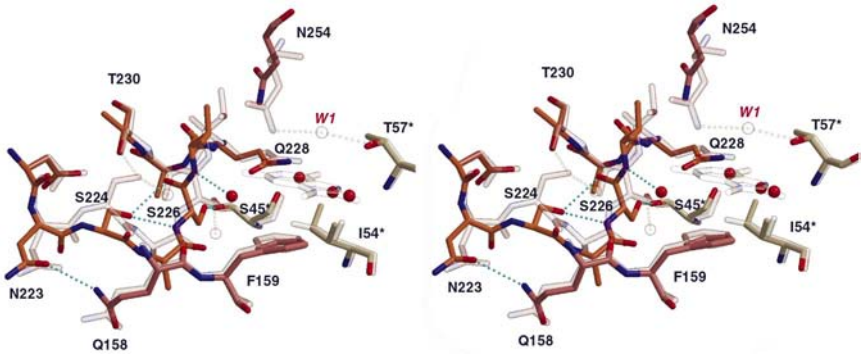

Figure 4
Stereoview of the superposition of the active-site region, highlighting the main differences between the ligand-free structure and complex (transparency) and the new hydrogen-bonding scheme (cyan dashed lines).

3.2. Overall structure

Uox is an active globular homotetrameric enzyme with an external size of 50 × 50 Å, enclosing a tunnel 50 Å long and 12 Å in diameter (Figs. 5*c* and 5*d*). The monomer is composed of two structurally equivalent domains, named tunnelling-fold (T-fold) domains (Colloc'h *et al.*, 1997). Each

domain displays a similar antiparallel superfold $\beta\beta\alpha\alpha\beta\beta$ with a α+β sandwich topology (H1–H2 + S1–S4 and H3–H4 + S5–S8, respectively). Additionally, they both contain an extra one-turn helix (h1 and h2). The two domains are joined by a four-residue turn between S4 and S5 (Fig. 5*a*). The asymmetry of domain 2 with respect to domain 1 is provided by the S7′–S8′ loop (260–277), which extends toward the outside with a linker of 18 residues, and a long coiled region, residues 153–178 (S5–S6 loop). The concatenation of the two T-fold domains gives rise to an antiparallel β-sheet consisting of eight long

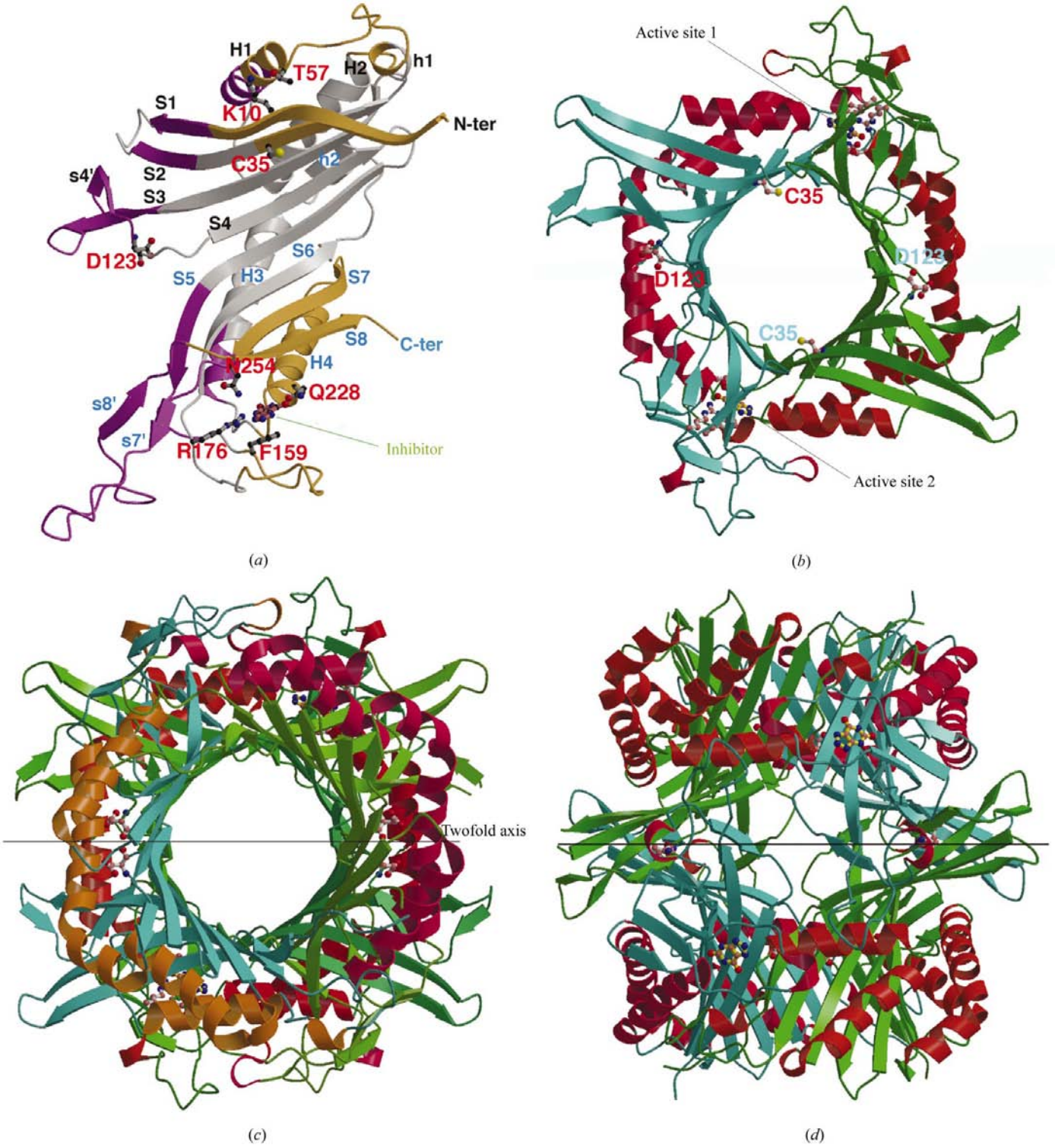

Figure 5
Urate oxidase topology. (*a*) The monomer (or asymmetric unit in the *I*222 space group); the residues involved in the tetrameric interface are shown in magenta and those involved in the dimeric interface in gold. (*b*) The dimer and (*c*) and (*d*) the tetramer (generated by the crystallographic 222 symmetries; two orthogonal views). Graphics created using *MOLSCRIPT* (Kraulis, 1991) and rendered using *Raster3D* (Merritt & Murphy, 1994).

sequential strands with the four α-helices layered on the concave side of the sheet. The upper parts of the strands appear to be involved in dimer interface and tunnel formation, whereas the lower parts appear to be involved in the tetrameric interface. Monomers *A* and *B* (and *C* and *D*) are related by a crystallographic twofold axis and are assembled as a dimer *via* perfect hydrogen bonding (more than 40 hydrogen bonds covering an interface-accessible surface area of $\sim$3700 Å^2, corresponding to $\sim$20% of the total surface) mainly between the upper parts of S1 and S8. The resulting barrel consists of an antiparallel β-sheet of 16 strands, with eight helices forming the outer part of the barrel (Fig. 5). The functional tetramer results from the union of two dimers ($A + B$ and $C + D$) stacked face-to-face through the lower part of the strands, with hydrogen bonds between strands S1 and S8 from *A* or *B* and strands S4 and S5 from *C* and *D* (Figs. 5*c* and 5*d*). The interface-accessible surface area between *A* and *C* (and *B* and *D*) is about 3100 Å^2, involving $\sim$26 hydrogen bonds. The crystallographic twofold axis which relates $A + B$ with $C + D$ passes between the symmetry-related Asp123 residues.

Proteins which belong to the T-fold family display some similarities in their active site, even though their functions differ. They all bind a substrate that possesses a pyrimidine or a purine ring. The relationship described between the proteins in the T-fold family (Colloc'h *et al.*, 2000) is still valid, even if the reversed position of AZA in urate oxidase modifies the superposition of the different active sites. A glutamate residue and the main-chain N atom of the preceding residue hold the pyrimidine ring of the substrate in 6-pyruvol tetrahydropterin synthase (PTPS), in dihydroneopterin aldolase (DHNA) and in GTP cyclohydrolase (CHY). Similarly, a glutamine (instead of a glutamate) and the N atom of the preceding residue hold the purine ring in urate oxidase. This conserved glutamate or glutamine residue is located at the N-terminal end of the second helix in the first monomer in all cases. A hydrophobic residue lies parallel to the pyrimidine ring in the PTPS, DHNA and CHY enzymes. In Uox, there is a threonine linked to the N7 atom of the inhibitor instead. This conserved hydrophobic (or threonine in Uox) residue is located at the N-terminal end of the first helix of the second monomer.

The urate oxidase sequences share 30–50% identity. This level of sequence identity allows the determination of which residues are likely to be conserved in the whole family (Fig. 6). 24 residues are conserved and of these seven are in topo-

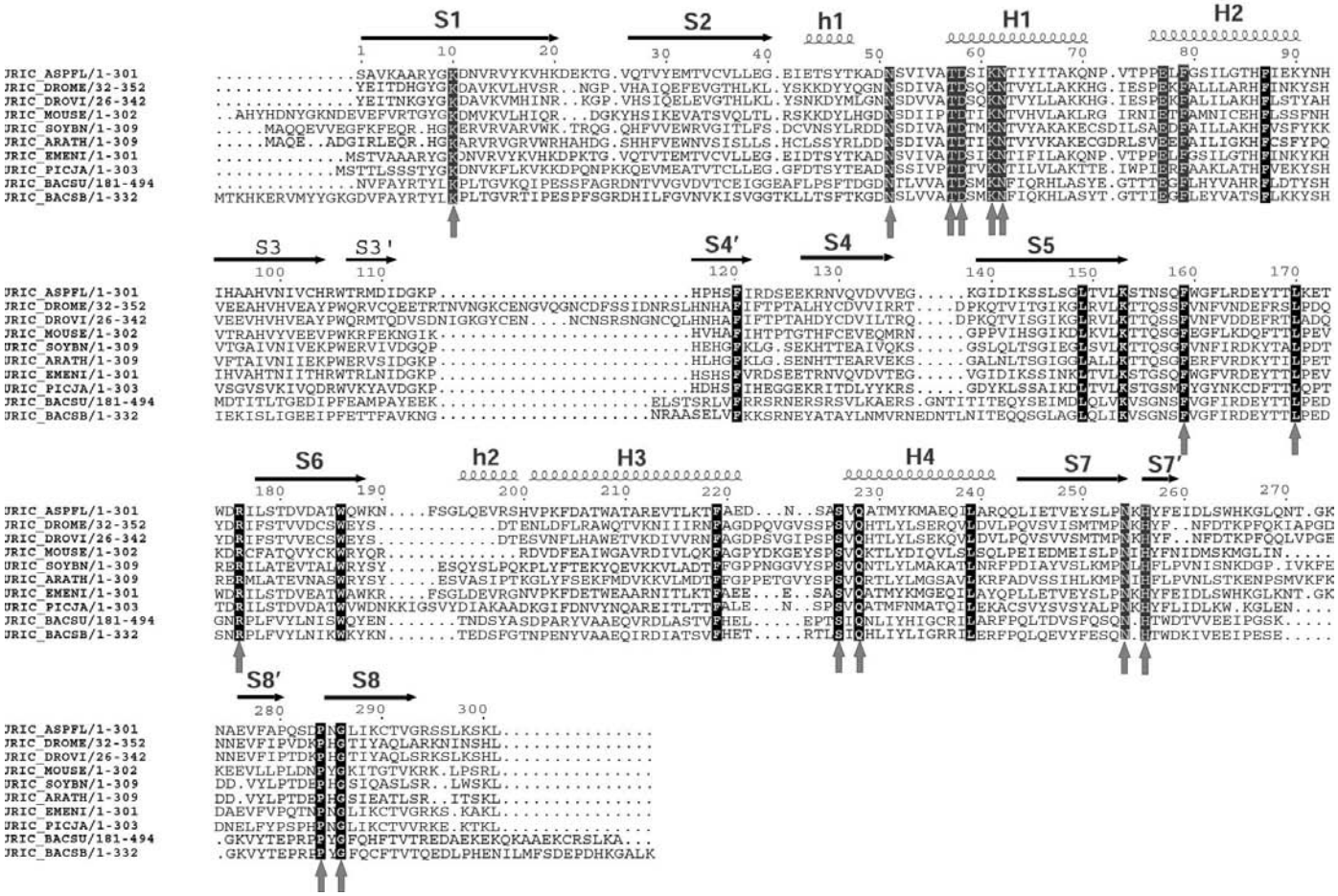

Figure 6
Alignment of ten remote sequences of the urate oxidase family, with their SWISS-PROT identifiers. The 24 conserved residues are shown in white on a black background. Of these, the 15 conserved residues which are close to the active site are indicated with arrows. The numbering and the secondary structure refers to the *A. flavus* urate oxidase. Helices and strands are named according to Colloc'h *et al.* (1997). (Produced using *ESPRIPT*; Gouet *et al.*, 1999.)

hydrophobic positions (Poupon & Mornon, 1998). It is interesting to note that among the conserved residues, there are no histidines nor cysteines, two amino acids that are often involved in oxidation processes. 15 conserved residues are close to the active site and may have a role in the function of the enzyme. The roles of Lys10, Thr57, Arg176, Gln228, Asn254 and His256 are investigated below. The Lys61 and Asp58 side chains are both hydrogen bonded to water molecules located in the entrance of the cavity and are linked together by a weak salt bridge (the distance between Lys61 N^{ζ} and Asp58 $O^{\delta1}$ is 3.13 Å). The side chain of Asn62 is too distant to be hydrogen bonded to Asp58. Two particular residues, Pro284 and Gly286, located between Lys10 and Asn254 may play a structural role in the conformation of the active site. In addition, Leu170 and Phe159 constitute a hydrophobic floor below the substrate; the leucine is located at the entrance cavity and may help in the positioning of the substrate. Finally, the two conserved residues Asn51 and Ser226 have no clear role in the function of the enzyme.

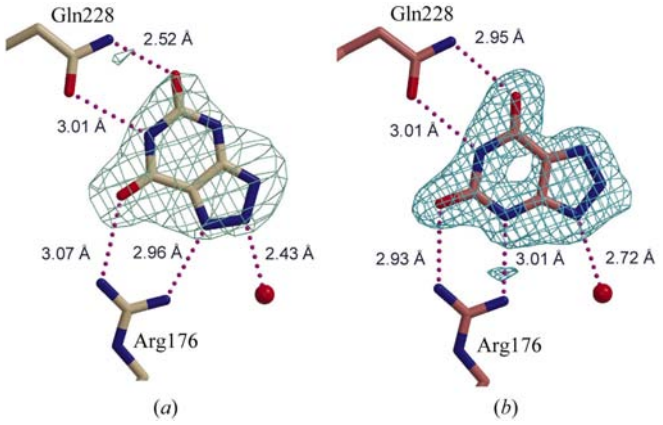

Figure 7
Revised position of 8-azaxanthin relative to the active site of Uox. (*a*) Original published location of the AZA inhibitor in 1uox; (*b*) the revised orientation of AZA in the new refined structure (1.8 Å). Both orientations are superimposed on the respective final $2|F_{obs}| - |F_{calc}|$ electron-density maps contoured around the inhibitor at 1.75σ.

3.3. Bad contacts between the carbonyl group of Asp123 and its twofold-symmetry-related counterpart

The asymmetric unit region in Fig. 2 was selected to present one problem of map interpretation: the carbonyl of Asp123 was obviously face-to-face with its twofold-related counterpart at a distance of 1.5 Å, since the C^{α} chain and neighbours appeared to fit the density appropriately. In the previous 1uox model, the carbonyls were not rigorously aligned but the distance between the two O atoms was rather too short (∼2.2 Å). Such a modelling was also produced (i) in an unbiased way with a model built automatically in a solvent-flattened xenon SIR electron-density map calculated elsewhere (work to be published) without prior knowledge of the molecular structure and (ii) in the four independent contact zones in the native $P2_1$ structure and other structures refined subsequently. This clearly demonstrates that that problem does not arise from any twofold-axis artifact.

Collision was avoided by an alternate flip of the peptide bond and refinement as a 1:1 disorder to comply with the strict crystallographic symmetry of the dimer and to promote CO—NH hydrogen bonds (see close-up in Fig. 2*b*).

3.4. Description of the active site

The homotetrameric enzyme appears to be symmetrically liganded with one inhibitor per active site (full occupancy). Each active site is formed at the bottom of a funnel-shaped pocket whose opening is created at the dimer interface by the flexible protuberant linker between the two strands S7′ and S8′, the coiled region 150–175 and their symmetrical counterparts. The same entrance is therefore used by two molecules per dimeric subunit, suggesting possible cooperativity. The inhibitor deeply penetrates into the crevice closed by the N-terminal region of H4 of one monomer and flanked by the three β-strands S6–S8 on one side and S1 and the N-terminal region of H1 from the second monomer on the other side. Calculation of the electrostatic potential using *GRASP* (Nicholls *et al.*, 1991) clearly shows that the cavity is positively charged and therefore likely to bind anionic species.

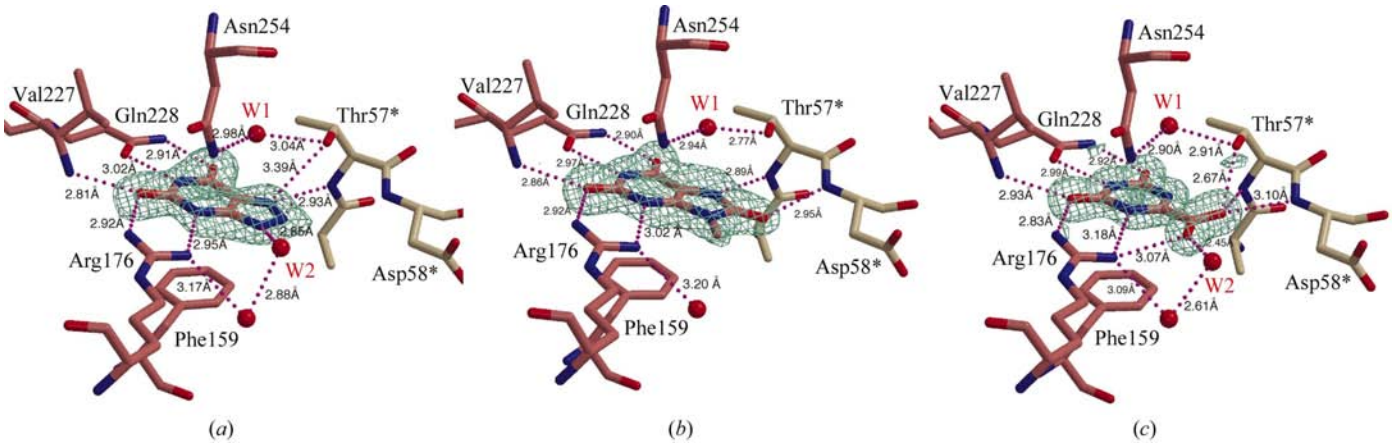

Figure 8
Omit maps of the three inhibitors: (*a*) AZA, (*b*) MUA, (*c*) OXC. Important water molecules are also shown, as well as distances to the closest residue side chains.

The improved map quality permitted us to reassign the position of the inhibitor 8-azaxanthin (Fig. 7). Compared with the initial orientation, inversion of the inhibitor plane leads to a completely different hydrogen-bonding scheme that is now more regular in terms of geometry with the neighbouring water molecules.

Additional confirmation was provided by the similar interactions between the six-membered rings of MUA and OXC with the protein-residue side chains (Fig. 8).

The two important residues Gln228 and Arg176 develop hydrogen-bond interactions and build molecular tweezers that handle the inhibitor in the three complexes. In addition, the inhibitor ring is stacked over the side chain of Phe159 and displays interactions with the main-chain N atom of Thr57* (where * represents monomer *B*); water-mediated interactions with Asn254 and Thr57* side chains are also observed. In MUA and OXC, additional interactions are present with the Asp58* main chain.

3.5. The water molecules in the active site

Apart from the ligand-free structure, the networks of water molecules in the active site are similar in the three complexes. A water molecule (W1 in Fig. 8) that is strongly suspected to be involved in the reaction is located above the C4—C5 bond in AZA or MUA (the C4—N5 bond in OXC) in an ideal position for hydroxylation. W1 is tightly hydrogen bonded to the side chains of Asn254 from one monomer and of Thr57* from a second monomer. In addition, Lys10* from a second monomer, which is known to be a crucial residue for the catalytic activity, is implicated in the stability of this hydrogen-bonding net by connecting Thr57* to His256, in agreement with Imhoff *et al.* (2003).

Another important water molecule, W2, is hydrogen bonded to N9 in AZA or to $O^{\delta 2}$ in OXC, but does not exist in MUA because of the methyl group bound to N9. W2 is hydrogen bonded to an extended network of nine water molecules which lies within the first monomer. This network is conserved in all the three complex structures.

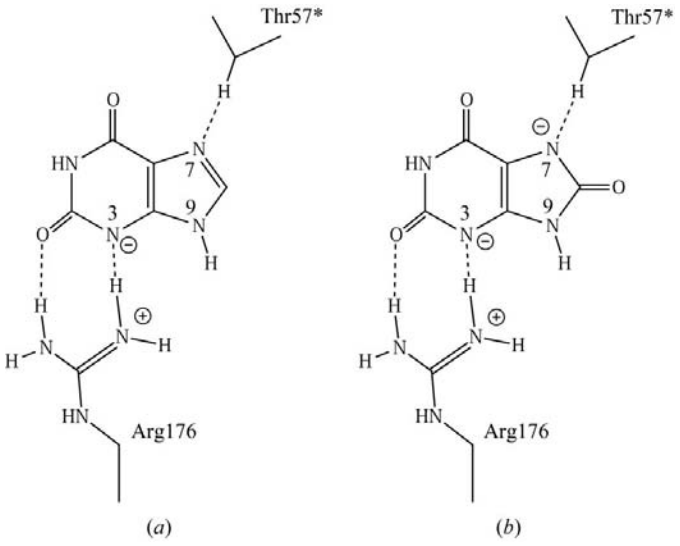

Figure 9
(*a*) Assignment of tautomeric protons for 8AZA complex; (*b*) the 3,7-dianion in the catalytic site.

3.6. Comparisons between bound and free states

In the absence of inhibitor, the new crystal structure of Uox revealed a slight expansion of the pocket owing to the backward motion of the region 223–228 and the side chain of Asn254 as the only significant structural differences (Fig. 4). As a consequence, the W1 molecule is observed in only three of the eight independent active sites present in the asymmetric unit.

At the level of atomic interactions, the new subsequent orientations of Ser224 and Ser226 provide visible rearrangements in the intramolecular hydrogen-bond network which is extended to the dimer interface *via* a new hydrogen bond between Ser226 O^{γ} and the Ser45* carbonyl. Rotation of the methyl-group side chain of Ile54 away from the phenyl ring of Phe159 together with the new hydogen bond between the side chains of Gln158 and Asn223 might appear as a direct consequence of the free state of the enzyme. Indeed, similar situations occur within the eight independent active sites present in the $P2_1$ asymmetric unit.

4. Mechanistic implications

Since uric acid is structurally close to 9-methyl uric acid or 8-azaxanthin, it is possible to build a model of the active site with the natural substrate positioned similarly to the inhibitors. A number of points are noteworthy in this model.

(i) The distance between N3 and NH1 of Arg176 (∼3 Å) allows room for one proton (Figs. 8 and 9). Similarly, the distance between O2 (bound to C2) and NH2 of Arg176 allows room for only one other proton. Assuming that Arg176 is normally protonated, this implies that uric acid is ionized in position 3.

(ii) The distance between N7 and the main-chain N atom of Thr57 (∼2.9 Å) also allows room for one proton. This proton necessarily belongs to the peptide bond. Hence, uric acid cannot bear an in-plane proton at position 7.

(iii) Experiments with *Bacillus subtilis* Uox have shown that Thr69 and Lys9 (which correspond to Thr57 and Lys10 in *A. flavus* Uox) are essential for activity, since the mutants T69A or K9M are inactive (Imhoff *et al.*, 2003).

The two first points imply that uric acid is bound to Uox as the 3,7-dianion (Fig. 9). Of all the putative dianions of uric acid, the 3,9-dianion has been shown to be the most stable when considered free in solution (Tipton, 2002). However, its introduction within the catalytic site implies a collision between the hydrogen borne by N7 and the main-chain amide group of Thr57. It can be assumed that in this situation the 3,9-dianion becomes less favourable than the 3,7-dianion. The 3,7-dianion may then have its H9 atom directly accessible, or indirectly accessible *via* the W2 molecule, to an incoming dioxygen molecule (as long as dioxygen uses this channel to enter the active site). Ongoing quantum-mechanical calculations suggest strong repulsion between H9 and W2 (Monard *et*

al., 2004). The very first step of the catalyzed reaction might therefore consist of a radical abstraction of the hydrogen from N9.

The third point is related to the hydroxylation process itself that leads to the primary intermediate hydroxylated at C5. It is likely that the W1 molecule, accurately identified in the complexed structures above the double bond C4═C5 and bridging the side chains of Asn254 and Thr57* (the latter interacting with another catalytically important residue, namely Lys10), is used by the enzyme to give the intermediate product as previously suggested (Colloc'h *et al.*, 1997).

5. Conclusion

The availability of higher resolution data from Uox-complex crystals and a reassessment of the structure have led to a revision of ligand orientation in the active site. The resulting new hydrogen-bonding scheme has profound implications for the proposed mechanism of catalysis. The structure of the free enzyme was also analyzed but in a different space group in which the full tetramer is retained. Furthermore, owing to packing considerations, the C-terminal end that interacts with a neighbouring tetramer is now stabilized and becomes visible in the electron density. Several different inhibitors were also explored: not only 8-azaxanthin but also 9-methyl uric acid and oxonic acid. They are oriented in a similar way to 8-azaxanthin. As long as they correctly mimic the uric acid orientation, they favour a mechanistic model starting with the abstraction by molecular oxygen of a proton that should be located on the N9 atom. Theoretical studies are under way in order to investigate the reaction pathway based on this hypothesis, as well as the subsequent steps of the reaction.

This work was partly supported by European commission grant No. HPRI-CT-1999-50015 within the EXMAD project. We thank Dr A. Bayol (Sanofi Corp.) for a gift of recombinant Uox, Dr J. Roach (UNC, Chapel Hill) and Dr A. E. Aplin (Albany Medical College) for proofreading the manuscript.

References

Ames, B. N., Cathcart, R., Schwiers, E. & Hochstein, P. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 6858–6862.
Bayol, A., Capdevielle, J., Malazzi, P., Buzy, A., Bonnet, M. C., Colloc'h, N., Mornon, J. P., Loyaux, D. & Ferrara, P. (2002). *Biotechnol. Appl. Biochem.* **36**, 21–31.
Bentley, R. & Neuberger, A. (1952). *Biochem. J.* **52**, 694–699.
Bonneté, F., Vivarès, D., Robert, C. & Colloc'h, N. (2001). *J. Cryst. Growth*, **232**, 330–339.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Colloc'h, N., El Hajji, M., Bachet, B., Lhermite, G., Schiltz, M., Prangé, T., Castro, B. & Mornon, J. P. (1997). *Nature Struct. Biol.* **4**, 947–952.
Colloc'h, N., Poupon, A. & Mornon, J. P. (2000). *Proteins*, **39**, 142–154.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Fourme, R., Dhez, P., Benoit, J. P., Kahn, R., Dubuisson, J. M., Besson, P. & Frouin, J. (1992). *Rev. Sci. Instrum.* **63**, 982–987.
French, G. S. & Wilson, K. S. (1978). *Acta Cryst.* A**34**, 517–525.
Gouet, P., Courcelle, E., Stuart, D. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Imhoff, R. D, Nicholas, P., Power, N. P., Borrok, N. J. & Tipton, P. A. (2003). *Biochemistry*, **42**, 4094–4100.
Irwin, J. & Bricogne, G. (1996). *Acta Cryst.* A**52**, C86.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kahn, K., Serfozo, P. & Tipton, P. A. (1997). *J. Am. Chem. Soc.* **119**, 5435–5442.
Kahn, K. & Tipton, P. A. (1997). *Biochemistry*, **36**, 4731–4738.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Kraulis, P. E. (1991). *J. Appl. Cryst.* **24**, 946–950.
Lamzin, V. & Wilson, K. S. (1997). *Methods Enzymol.* **277**, 269–305.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Modric, N., Derome, A. E., Ashcroft, S. J. H. & Poje, M. (1992). *Tetrahedron Lett.* **33**, 6691–6694.
Monard, G., Altarsha, M. & Castro, B. (2004). Submitted.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 306–326.
Perrakis, A., Morris, R. J. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Poupon, A. & Mornon, J. P. (1998). *Proteins*, **33**, 329–342.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
Raychaudhuri, A. & Tipton, P. A. (2002). *Plant. Physiol.* **130**, 2061–2068.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Roversi, P., Blanc, E., Vonrhein, C., Evans, G. & Bricogne, G. (2000). *Acta Cryst.* D**56**, 1316–1323.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–341.
Tipton, P. A. (2002). *Methods Enzymol.* **354**, 310–319.
Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
Vivarès, D. & Bonneté, F. (2002). *Acta Cryst.* D**58**, 472–479.
Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
Wu, X., Muzny, D. M., Lee, C. C. & Caskey, C. T. (1992). *J. Mol. Evol.* **34**, 78–84.