Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The crystallographic structure of the aldose reductase–IDD552 complex shows direct proton donation from tyrosine 48

Federico Ruiz,[a] Isabelle Hazemann,[a] Andre Mitschler,[a] Andrzejj Joachimiak,[b] Thomas Schneider,[c] Martin Karplus[d,e] and Alberto Podjarny[a]*

[a]UPR de Biologie Structurale, IGBMC, CNRS INSERM ULP, 1 Rue Laurent Fries, BP 163, 67404 Illkirch, France, [b]Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois, USA, [c]FIRC Institute of Molecular Oncology, Via Adamello 16, 20139 Milano, Italy, [d]Laboratoire de Chimie Biophysique, ISIS/ULP, 8 Allée Gaspard Monge, BP 70028, F-67083 Strasbourg CEDEX, France, and [e]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts, USA

Correspondence e-mail: podjarny@titus.u-strasbg.fr

Received 10 March 2004
Accepted 10 May 2004

PDB References: aldose reductase–IDD552 complex, pH 8, 1t41, r1t41sf; pH 5, 1t40, r1t40sf.

The X-ray crystal structure of human aldose reductase (ALR2) in complex with the inhibitor IDD552 was determined using crystals obtained from two crystallization conditions with different pH values (pH 5 and 8). In both structures the charged carboxylic head of the inhibitor binds to the active site, making hydrogen-bond interactions with His110 and Tyr48 and electrostatic interactions with $NADP^+$. There is an important difference between the two structures: the observation of a double conformation of the carboxylic acid moiety of the inhibitor at pH 8, with one water molecule interacting with the main configuration. This is the first time that a water molecule has been observed deep inside the ALR2 active site. Furthermore, in the configuration with the lower occupancy factor the difference electron-density map shows a clear peak (2.5σ) for the H atom in the hydrogen bond between the inhibitor's carboxylic acid and the Tyr48 side-chain O atom. The position of this peak implies that this H atom is shared between both O atoms, indicating possible direct proton transfer from this residue to the inhibitor. This fact agrees with the model of the catalytic mechanism, in which the proton is donated by the Tyr48 hydroxyl to the substrate. These observations are useful both in drug design and in understanding the ALR2 mechanism.

1. Introduction

Aldose reductase (ALR2; EC 1.1.1.21) is a monomeric member of the aldo–keto reductase superfamily. Members of this NADPH-dependent oxidoreductase superfamily are found in vertebrates, invertebrates, plants, protozoa, fungi, eubacteria and archaebacteria, implying that is one of the ancient superfamilies of enzymes (Jez & Penning, 2001). In the direct reaction (optimum pH 6; Schlegel *et al.*, 1998), ALR2 catalyzes the NADPH-dependent reduction of diverse aldehydes to their corresponding alcohols (Warren *et al.*, 1993; Jez *et al.*, 1997); for the inverse reaction the optimum pH is 9 (Schlegel *et al.*, 1998). ALR2 is the first and the rate-determining enzyme in the polyol pathway, an alternative route for glucose phosphorylation. Under hyperglycaemic conditions (*e.g.* diabetes), this pathway processes up to one-third of the total glucose turnover. The increased activity of ALR2 results in a decrease in the NADPH:$NADP^+$ ratio and impacts on other NADPH-dependent enzymes. Furthermore, accumulation of excess sorbitol has been linked with the development of diabetic complications such as retinopathy, neuropathy and nephropathy (Kinoshita & Nishimura, 1988). One of the possible options in the pharmaceutical development of compounds to act against these pathologies is to block

the activity of ALR2. Many aldose reductase inhibitors (ARIs) have been developed in recent years. Although only a few of them are now on the market, several new potential candidates are currently undergoing clinical trials (Pfeifer *et al.*, 1997; Oates & Mylari, 1999; Miyamoto, 2002).

The ARIs usually have a polar group attached to a hydrophobic ring system (Costantino *et al.*, 2000). They can be grouped into two main classes: carboxylic acid derivatives and cyclic imides (mostly hydantoins). The activities of both groups are similar *in vitro*, but cyclic imide derivatives are usually more potent *in vivo*. The substitution of an acetic acid group by a cyclic imide derivative has been shown to transform compounds active as ARIs almost exclusively *in vitro* to compounds with *in vivo* potency (Malamas & Hohman, 1994). With a pK_a value lower than four, the carboxylic acid derivatives are ionized at physiological pH, meaning that they have difficulty crossing biological membranes. In contrast, for cyclic imide derivatives the pK_a value is in the range 7–9.

The aldose reductase inhibitor IDD552 (Fig. 1) is a carboxylic acid derivative, as is the inhibitor IDD594 (which forms a complex that was solved at ultrahigh resolution; Howard *et al.*, 2004). The IC_{50} values of IDD552 against ALR2 and ALR1 are 11 n*M* and 22 µ*M*, respectively (M. van Zandt, personal communication).

ALR2 folds in an α/β-barrel structure with a core of eight parallel β-strands connected by peripheral α-helices (Fig. 2). Crystal structures of ALR2–inhibitor complexes have shown that the polar groups of the inhibitors bind in an anion-binding site between the nicotinamide ring of the coenzyme and the following ALR2 residues: Tyr48, His110 and Trp111 (Wilson *et al.*, 1993; Harrison *et al.*, 1997; Urzhumtsev *et al.*, 1997; Calderone *et al.*, 2000; Oka *et al.*, 2000; El-Kabbani *et al.*, 2004; Howard *et al.*, 2004). Some inhibitors induce a conformational change which opens a 'specificity' pocket (Urzhumtsev *et al.*, 1997) in which they interact with C-terminal residues that are not conserved in aldehyde reductase (ALR1). This other member of the aldo–keto reductase superfamily is closely related to ALR2 (65% identity). In order to test the inhibitor selectivity for ALR2, the activities of ARIs against ALR1 are usually assessed (Miyamoto, 2002).

The structures of aldose reductase complexed with IDD594 (Howard *et al.*, 2004), Fidarestat or Minalrestat (El-Kabbani *et al.*, 2004) have been solved at atomic or better (subatomic) resolution. Atomic resolution is reached when the crystallographic data have a resolution higher than 1.2 Å, with 50% of the theoretically measurable reflections in the outer shell above the noise level [$I > 2\sigma(I)$; Sheldrick, 1990]. The major advantage of structure determination at these high-resolution ranges is the greater ratio of experimental observations to parameters; the structure can therefore be refined using only weak restraints (Deacon *et al.*, 1997). The quality of electron-density maps makes it possible to identify the atom type and bond order and to model alternative residue conformations and sometimes to directly observe some of the H atoms (Dauter *et al.*, 1997; Longhi *et al.*, 1998; Schmidt & Lamzin, 2002).

In this work, we present the structure of human aldose reductase complexed with the inhibitor IDD552 obtained under two different crystallization conditions, pH 5 and pH 8, covering the whole range of enzymatic activity of this protein. Atomic resolution (1.0 Å) is achieved at pH 8. In this ALR2–IDD552 structure, the carboxylic head of the inhibitor presents a double conformation. Between one of the O atoms of this moiety and the Tyr48 side-chain O atom a clear peak is observed in a difference electron-density map, implying that the H atom is shared. This is the first crystallographic evidence of direct proton transfer from this residue to an inhibitor.

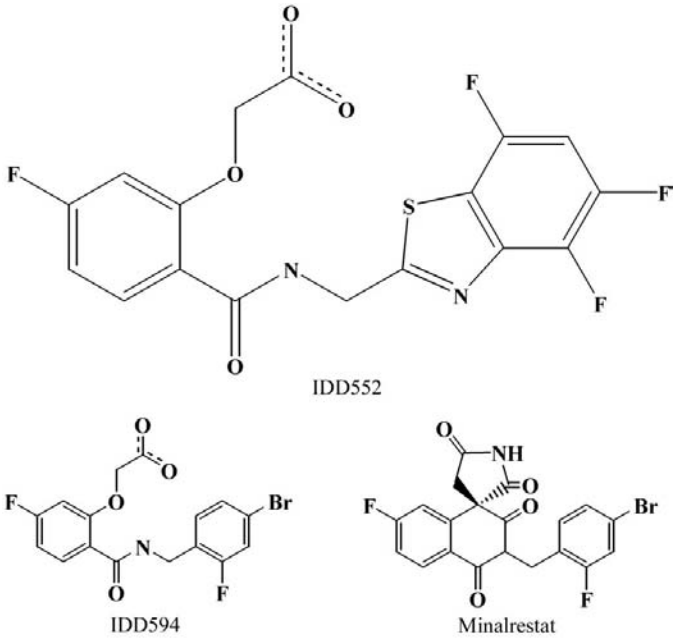

Figure 1
Chemical structures of the aldose reductase inhibitors IDD552, IDD594 and Minalrestat.

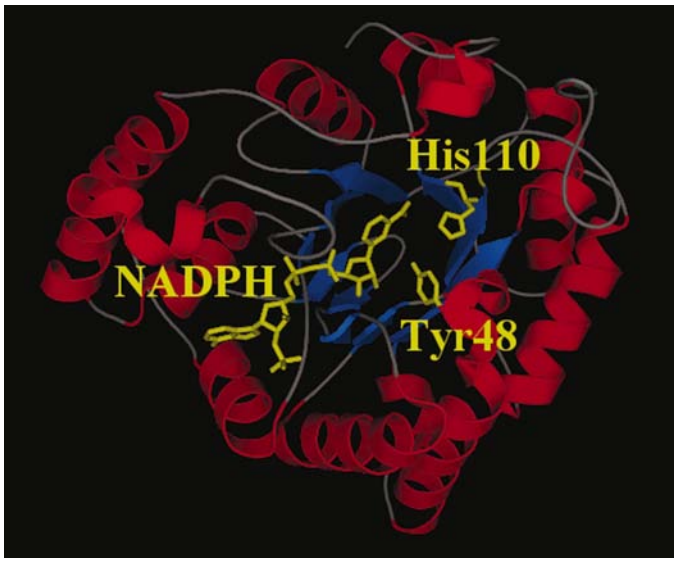

Figure 2
ALR2 α/β-barrel structure with α-helices in red, β-strands in blue and the NADPH cofactor and two active-site residues (His110 and Tyr48) in yellow.

Table 1
Data-collection statistics for aldose reductase–IDD552 complexes.

For the atomic resolution structure three different data sets were collected: a low-resolution data set ($d > 1.8$ Å) and two high-resolution data sets ($2.8 > d > 1.0$ Å). The intensities were merged at the end. Statistics from the low-resolution treatment are marked (l) and those from the highest resolution shell (1.00–1.04 Å) are marked (h). In the 1.8 Å structure the highest resolution shell (h) is 1.8–1.86 Å.

	pH 8	pH 5
No. crystals used	1	1
Wavelength (Å)	0.90042	0.9797
Space group	*P*1	$P2_1$
Unit-cell parameters		
a (Å)	47.06	50.00
b (Å)	47.52	67.24
c (Å)	40.03	47.78
α (°)	67.3	90
β (°)	76.8	92.5
γ (°)	75.8	90
Diffraction data		
Resolution range (Å)	50.0–1.0	16.0– 1.8
Unique reflections	151945	29192
$R(I)_{\rm sym}$ (l)	3.8	—
$R(I)_{\rm merge}$ (overall) (%)	—	4.7
$R(I)_{\rm sym}$ (h) (%)	35.8	17.0
Completeness (l) (%)	91.6	—
Completeness (overall) (%)	—	95.5
Completeness (h) (%)	73.7	80.0
Redundancy (l)	4.9	—
Redundancy (overall)	—	1.5
Redundancy (h)	3.6	1.4
$I/\sigma(I)$ (l)	20.4	—
$I/\sigma(I)$ (overall)	—	1.85
$I/\sigma(I)$ (h)	12.0	5.4

2. Materials and methods

2.1. Expression and purification

The open reading frame of the human aldose reductase gene (Accession GenBank/EMBL Data Bank No. J055017) was amplified by PCR from cDNA and cloned into the T7 RNA polymerase-based vector pET15b (Novagen). The protein was expressed in *Escherichia coli* strain BL21(DE3) (Novagen) induced by IPTG (Euromedex) during a 3 h period at 310 K. The pellet from a 6 l culture was disrupted by sonication and centrifuged. The supernatant was loaded onto a Talon metal-affinity column (Clontech). After thrombin cleavage of the hexahistidine extension, the protein was finally loaded onto a DEAE-Sephadex A-50 column (Pharmacia) and eluted with an NaCl gradient (Lamour *et al.*, 1999).

2.2. Crystallization

2.2.1. pH 8. Prior to crystallization, human ALR2 at 30 mg ml^{-1} in Tris–HCl buffer pH 8 was mixed with NADP$^+$ and inhibitor (the molar ratio of ALR2:NADP$^+$:inhibitor was 1:2:2). The ternary complexes were crystallized using the hanging-drop method (Linbro 24-well culture plates, Flow Laboratories). The ALR2/NADP$^+$/inhibitor solution was mixed with an equal volume of well solution (1 ml) containing 20% PEG 6000 and 0.1 *M* Tris–HCl. Crystals were grown at 277 K, transferred into a stabilization solution (25% PEG 6000) and then into cryoprotectant solution (40% PEG 6000) and dipped into either liquid nitrogen or ethane (Hazemann *et al.*, in preparation).

2.2.2. pH 5. The hanging-drop method was used to obtain crystals (Linbro 24-well culture plates, Flow Laboratories). Protein was co-crystallized with NADP$^+$(Sigma) and inhibitor at room temperature (protein:coenzyme:inhibitor ratio = 1:2:2). Previously equilibrated (ammonium citrate buffer, 15% PEG 6000) hanging drops were seeded with stock seed solutions diluted 100 times (Lamour *et al.*, 1999). Cryofreezing was carried out through rapid transfer into a stabilization solution (25% PEG 6000) and then into a cryoprotecting solution (40% PEG 6000) and finally into either liquid nitrogen or ethane.

2.3. Data collection and processing

2.3.1. pH 8. These experiments were carried out on X-ray beamline 19ID of SBC-CAT at the APS. The X-ray wavelength used was 0.90042 Å. Data from the crystals were reduced and scaled with *HKL*2000 (Otwinowski & Minor, 1997). Crystals belong to space group *P*1, with unit-cell parameters $a = 47.06$, $b = 47.52$, $c = 40.03$ Å, $\alpha = 67.3$, $\beta = 76.8$, $\gamma = 75.8^\circ$ and one complex per asymmetric unity. The resolution range is 1.0–50 Å, with 151 945 unique reflections.

2.3.2. pH 5. Crystallographic data were collected at the BM30 beamline at the ESRF (Table 1). Data from the crystals were reduced and scaled with *HKL*2000 (Otwinowski & Minor, 1997). Crystals belong to space group $P2_1$, with unit-cell parameters $a = 50.00$, $b = 67.24$, $c = 47.78$ Å, $\beta = 92.5^\circ$. The resolution range is 1.8–16 Å, with 29 192 unique reflections.

2.4. Structure refinement

For both complexes, the coordinates of human ALR2 holoenzyme were used to solve the structure by molecular replacement using the *CNS* program (Brünger *et al.*, 1998). The 1.8 Å structure was completely refined using this program.

For the atomic resolution structure, anisotropic conjugate-gradient refinement was carried out using the *SHELX* program package (Sheldrick & Schneider, 1997). Restraints were applied to bond lengths and angles, chiral volumes, planarity of aromatic rings and van der Waals contacts ('anti-bumping restraints'). In the final cycles, H atoms were introduced. Water molecules were located in a difference map. The model was manually corrected using $(2F_o - F_c)$ and $(F_o - F_c)$ electron-density maps displayed with the *XtalView*/*Xfit* program (McRee, 1999). Standard uncertainties in atomic positions and bond lengths were calculated from a full-matrix inversion. Refinement data are shown in Table 2. Structural figures were produced using *PyMol* (DeLano, 2002).

2.5. pK_a calculation

The atomic resolution structure of ALR2–IDD552 was used in the pK_a calculations. The water molecules were deleted except for those near the active site (five molecules). Inclusion of water molecules in the vicinity of the region of interest is necessary to fill cavities in the structure which could (incorrectly) be assigned a high dielectric constant corresponding to

Table 2
Refinement statistics for the aldose reductase–IDD552 complexes.

	pH 8		pH 5
Refinement	Anisotropic ($F > 4\sigma$)	Anisotropic (all reflections)	Isotropic
Resolution range (Å)	10.0–1.0		8.0–1.8
Reflections used (working/test set)	103682/3198	130180/4050	25751/2133
R_{cryst}/R_{free} (%)	—	—	15.9/19.5
With H	9.8/12.2	10.8/13.3	—
Without H	10.9/13.1	12.0/14.4	—
No. protein residues (mean *B* factor, Å^2)	316 (13 ± 9)	316 (15 ± 8)	
No. coenzyme molecules (mean *B* factor, Å^2)	1 (7 ± 2)	—	1 (9 ± 2)
No. inhibitor molecules (mean *B* factor, Å^2)	1 (8 ± 1)	—	1 (13 ± 1)
No. water molecules (mean *B* factor, Å^2)	338 (26 ± 1)	—	263 (30 ± 1)
Ramachandran plot			
Residues in most favourable region (%)	91.4		91.3
Residues in additional allowed regions (%)	8.6		8.7

bulk water (Gibas & Subramaniam, 1996). Atomic radii and partial atomic charges were assigned according to the *CHARMM* all-hydrogen parameter set 22 (McKerell *et al.*, 1998), $NADP^+$ parameters are taken from the literature (Pavelites *et al.*, 1997). The parameters for the inhibitor were obtained using *QUANTA* (Accelrys Inc.). The IDD552 atomic charges were obtained using the 6-31G++ basis set in *Gaussian*98 (Frisch *et al.*, 1998). The inhibitor was considered to be charged, *i.e.* as an ionized carboxyl group. The inhibitor was not included in pK_a calculations as a titrable residue.

The linearized Poisson–Boltzmann equation (Fogolari *et al.*, 2002) was numerically solved using the *UHBD* program (Davis *et al.*, 1991). The dielectric constants were set to 80 and 20 for the solvent region and protein interior, respectively. The ionic strength used in the calculations was 145 m*M* (physiological strength). The temperature was 273 K and a Stern (ion-exclusion) layer of 2 Å was used. A previously developed (van Vlijmen, Curry *et al.* 1998) shell script (Michael Schaefer, unpublished) was employed to perform the calculations.

The adequacy of a single crystallographic structure for pK_a determination has been examined and it was suggested that better results are obtained if an average over a set of molecular-dynamics structures is used (van Vlijmen, Schahefer *et al.*, 1998). However, crystallographic structures have been used successfully to evaluate the pK_a of residues in the active site of hen egg-white lysozyme, excluding structures with crystal contacts present near this site (Nielsen & McCammon, 2003). The authors argued that the errors produced by the crystallographic structure in the pK_a-value determination arise from the geometry of the surface residues imposed by crystallographic contacts. The active-site residues of ALR2 are far from the protein surface, allowing pK_a-value determination directly from the crystallographic structure. In accordance with this hypothesis, a medium-resolution structure has been used to identify the ALR2 active-site residues from the titration curves (Ondrechen *et al.*, 2001).

2.6. Molecular-dynamics study

The molecular-dynamics study was performed using the *CHARMM*-22 program (Brooks *et al.*, 1983).

A 30 Å radius sphere filled with TIP3P water molecules (Jorgensen *et al.*, 1983) was overlaid on the inhibitor-binding site (centred at the C atom of the IDD552 carboxylic acid moiety). Water molecules that overlap atoms of the complex (minimum distance 2.8 Å) or are beyond the 30 Å radius (including the crystallographic water molecules) were eliminated.

Weak harmonic forces restrained the atoms in the external region (those between 25 and 30 Å from the centre of the sphere) to their initial positions. Atoms of the protein outside the 30 Å sphere were fixed.

The atomic charges of titrable side chains in the 'buffer' region were reduced by scaling factors to mimic the solvent screening (Simonson *et al.*, 1997; Zeng *et al.*, 1999). The scaling factor is the ratio between the electrostatic potentials calculated using the Poisson–Boltzman equation with the 30 Å sphere surrounded by vacuum or by a high dielectric solvent. With this correction, the potential in the centre of the sphere accurately reproduces the potential that would be observed in bulk solvent, even when it is surrounded by vacuum. The charges of ionized residues outside the 30 Å were scaled by 80, the dielectric constant of water.

The molecular-dymanics simulations were carried out at a constant temperature of 300 K, with a 2 fs time step and stochastic boundary conditions (Leach, 2001).

3. Results

3.1. pH 8

The final model includes 316 residues, $NADP^+$, IDD552 and 338 solvent molecules. Multiple conformations were observed for 52 residues and 27 water molecules have double sites. 13 protein residues present at least one of their atoms in a single conformation with occupancy lower than one (partially occupied sites).

In a Ramachandran plot, 99.7% of the residues (excluding Gly) are in the core region (all residues except Asn50, $\varphi = -140.6$, $\psi = 12.6^\circ$) and 83.6% (excluding Gly) are in the inner core region (Kleywegt & Jones, 1996)

3.2. pH 5

The final model includes 316 residues, $NADP^+$, IDD552 and 263 solvent molecules. In a Ramachandran plot 91.3% of the residues are in the most favoured region and 8.7% are in the additional allowed region. Only three residues (Lys21, Asp134 and Pro222) are not in the core region (these disordered residues are in the surface of the protein).

The stereochemistry of both structures was checked using *PROCHECK* v.3.4.4 (Morris *et al.*, 1992).

3.3. Atomic resolution features of the ALR2–IDD552 complex structure at pH 8

The high quality of the crystallographic data is reflected in the details of the electron-density maps. The individual atoms can be distinguished. In ordered regions, the shape of the density at single and double bonds is different. For example, at residue Asp105 the electron density has continuity between the C^{γ} atom and one of the O atoms, indicating a double bond between these atoms, as has been observed in other proteins at atomic resolution (Minasov *et al.*, 2002).

In ($F_o - F_c$) H-omit maps some H atoms are seen attached to main-chain atoms and even to side-chain atoms. This fact, together with the flatness of the nicotinamide ring, allows us to conclude that the cofactor is $NADP^+$.

The structure reveals 52 double conformations. It is not possible to distinguish these double conformations at 1.8 Å resolution, suggesting that low-resolution models fit only one of several conformations, giving a wrong final model with a high-energy geometry (Minasov *et al.*, 2002).

The major deviation from planarity of an aromatic ring was observed in Tyr209 (r.m.s. from planarity, 0.046 Å; $\langle B \rangle$ for C atoms = 5.7 $Å^2$). This residue makes van der Waals contacts with the nicotinamide ring of $NADP^+$.

Some differences from the standard chemical values have been observed. 17 peptide bonds deviate by more than 2σ ($\sigma = 5.8^\circ$) from the 180° standard value. An example is the ω angle value for Lys77: -167.3° (a deviation from the standard value of larger than 10°, as observed in the ultrahigh-resolution structure; Howard *et al.*, 2004). This geometry is stabilized by a hydrogen bond between the N atom of Lys77 and the O^{γ} atom of Ser76. Such deviations from the ideal value are not uncommon and have been found throughout the PDB, especially in high-resolution structures (Ridder *et al.*, 1999).

3.4. Inhibitor binding at pH 8 and comparison with the ALR2–IDD552 structure at pH 5

Even when the crystallization conditions and the space group are different, the complex geometry is nearly the same in both cases. The maximum r.m.s. differences are observed in the side chains of several residues located on the protein surface. These differences could be linked to different protonation states (not proven at 1.8 Å) of these residues (establishing different electrostatic interactions) or to different interactions with symmetry-related molecules. The main difference between the two structures concerns the inhibitor polar-head geometry.

The final ($2F_o - F_c$) α_A-weighted map (Fig. 3) clearly shows the atomic electron density in the inhibitor-binding site. His110 and Tyr48 make hydrogen bonds with the inhibitor. The $F_o - F_c$ difference electron-density map (Fig. 4) shows that His110 is neutral and is singly protonated at the $N^{\varepsilon 2}$ position. After a refinement cycle without bond-length restraints, the $C^{\varepsilon 1}$—$N^{\delta 1}$ bond length (1.36 ± 0.01 Å) is shorter than the $C^{\varepsilon 1}$—$N^{\varepsilon 2}$ bond length (1.39 ± 0.01 Å). All these observations indicate that His110 is singly protonated at the $N^{\varepsilon 2}$ position as in previously determined structures (Howard *et al.*, 2004).

The 1.0 Å resolution clearly shows all the atoms of the inhibitor, including a double conformation of the carboxylic acid moiety of IDD552 (Fig. 4). The characteristics of these two conformations are as follows.

(i) The main conformation (occupancy 63.6%) is reported here for the first time. In this configuration, the O34 atom is placed at 2.47 ± 0.01 Å from Tyr48 OH and at 2.84 ± 0.01 Å from His110 $N^{\varepsilon 2}$. In this configuration the inhibitor does not interact directly with Trp111; a water molecule acts as water bridge between IDD552 and this last residue (the water

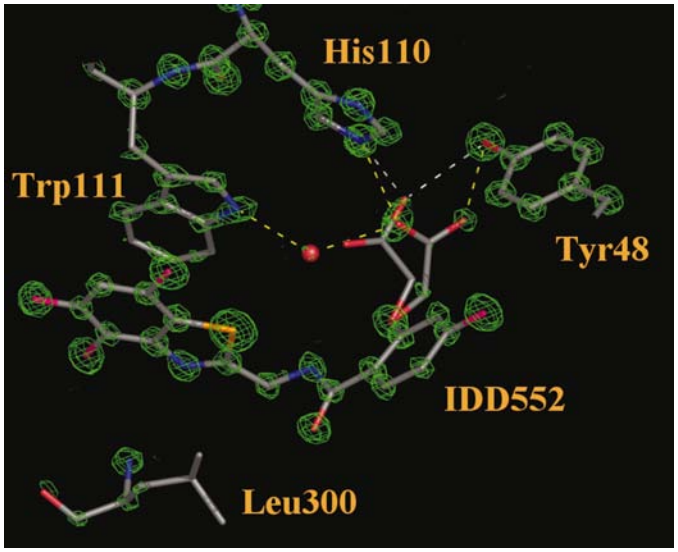

Figure 3
Inhibitor-binding site observed at 1.0 Å resolution. A ($2F_o - F_c$) electron-density map contoured at 4σ is shown in green. At this σ level no clear density is observed around the carboxylic acid moiety of the inhibitor owing to the double conformation.

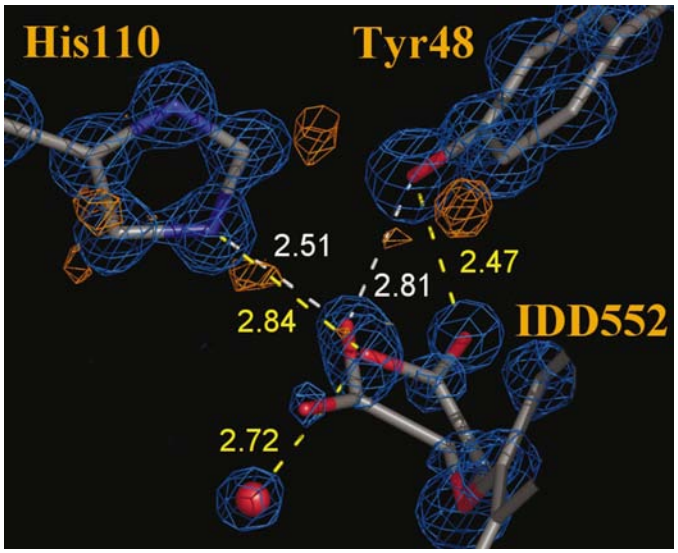

Figure 4
Double conformation of the IDD552 carboxylic acid moiety at pH 8. A ($2F_o - F_c$) electron-density map contoured at 2σ is shown in marine and a ($F_o - F_c$) electron-density map contoured at 2σ is shown in orange. Distances in Å (in yellow for the conformation with higher occupancy).

molecule is placed at 2.72 ± 0.01 Å from the O33 atom of the carboxylic acid moiety and at 2.77 ± 0.01 Å from Trp111 $N^{\varepsilon 1}$). This is the first time that a water molecule has been observed inside the active site of ALR2, a polar region on top of a hydrophobic cleft.

Furthermore, in this configuration the carboxylic acid moiety has a shorter distance, compared with previous measured structures, from the C4 atom of the nicotinamide ring of $NADP^+$ (Fig. 5). This last atom is the hydride donor in the enzymatic mechanism.

(ii) For the second conformation (occupancy 36.4%), a positive peak in an ($F_o - F_c$) electron-density map suggests that the H atom of the Tyr48 hydroxyl group is shared with the inhibitor (Fig. 6*b*). Similar distances from a peak to heavy atoms have been observed at high resolution, where an H atom is shared between the heteroatoms (Kuhn *et al.*, 1998). Note that in the previously reported AR–inhibitor structures the H atom of the Tyr48 hydroxyl is observed in a standard position (see Fig. 6).

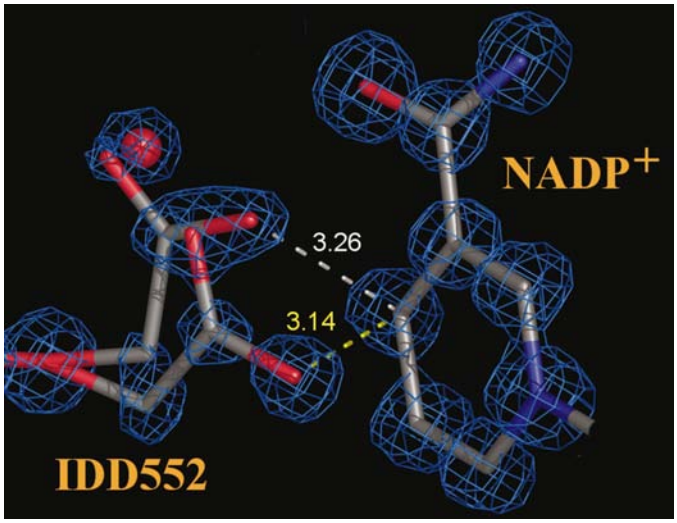

Figure 5
The interaction IDD552–$NADP^+$ in the atomic resolution structure. Distances in Å (in yellow for the conformation with higher occupancy).

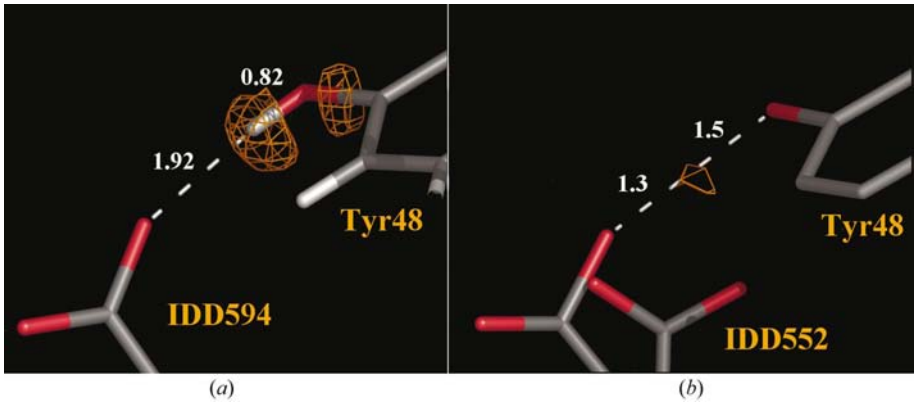

Figure 6
Comparison between the electron-density maps (contoured at 2σ in orange) around the O atom of the carboxylic head of the inhibitor. In the ALR2–IDD594 structure (*a*) the H atom is clearly bound to the O atom of Tyr48, a configuration different to that observed in the ALR2–IDD552 complex at atomic resolution (*b*). Note also the bond density only observed in the ALR2–IDD594 complex owing to the ultrahigh resolution (0.66 Å).

In this ALR2–IDD552 configuration, the O34 atom of the carboxylic moiety is placed at 2.81 ± 0.01 and 2.51 ± 0.01 Å from Tyr48 and His110, respectively. These distances are similar to those observed in previous determined ALR2 complex structures, in particular those obtained at lower resolution for the same complex (2.88 Å from Tyr48 and 2.71 Å from His110) or those of the ALR2–IDD594 complex (2.740 ± 0.005 Å from Tyr48 and 2.670 ± 0.005 Å from His110).

The inhibitor binds inside the 'specificity' pocket between Trp111 (the inhibitor is placed parallel to the aromatic ring of this residue, with a distance of 3.5 Å between the rings) and Leu300. The interactions inside this pocket are essentially hydrophobic. An interaction was observed between one of the F atoms of the inhibitor and the S atom of Cys80 (distance 3.34 ± 0.01 Å). This is the main difference from the binding of other inhibitors, such as IDD594 (Howard *et al.*, 2004) or Minalrestat (El-Kabbani *et al.*, 2004), which have an electrostatic interaction with Thr113 O^{γ}, a residue at the end of the 'specificity' pocket.

3.5. pK_a calculations

The position of the IDD552 carboxylic moiety produces changes in the pK_a values of the active-site residues. If the inhibitor is deleted (its volume is considered as a high dielectric region during the calculations) the pK_a value for His110 is 0.7 and that for Tyr48 is 9.6 (similar to the values of 0.9 and 8.5, respectively, that have been obtained in previous studies; Varnai & Warshel, 2000; Ondrechen *et al.*, 2001). The main message from these results for His110 is qualitative, in the sense that they confirm the hypothesis based on the holoenzyme crystallographic structure: the His110 neutral state is favoured (implying a lower pK_a value) because it is close to three hydrophobic residues (Trp79, Val47 and Trp111; Wilson *et al.*, 1992).

If the higher occupancy geometry is considered, the values are 1.8 and 12.8 for His110 and Tyr48, respectively. When the conformation with lower occupancy is considered, the values are 2.3 (for His110) and 11.1 (for Tyr48). The pK_a values of other titrable residues are not affected at all by the presence or geometry of the inhibitor.

3.6. Molecular dynamics

The dihedral angle involved in the double configuration of the inhibitor was analyzed using a 750 ps molecular-dynamics run. In the crystallographic structure, the values observed for this dihedral are 16 and −13.6° for the configurations with higher and lower occupancies, respectively. Starting with the first of these configurations, with the

associated water molecule, the value obtained is $9 \pm 9^{\circ}$. The distance between both O atoms of the inhibitor acid moiety and Tyr48 (2.7 ± 0.1 Å, crystallographic value 2.47 ± 0.01 Å), His110 (3.0 ± 0.2 Å, crystallographic value 2.84 ± 0.01 Å) and the water molecule (2.7 ± 0.1 Å, crystallographic value 2.72 ± 0.01 Å) were conserved. If the water molecule is eliminated, the dihedral angle changes and its value is reduced (mean value $4 \pm 12^{\circ}$), but does not reach the crystallographic configuration with lower occupancy.

4. Conclusions

Two ALR2–IDD552 complex structures (1.8 Å resolution, pH 5 and 1.0 Å resolution, pH 8) are presented in this work. As in the previous complexes determined at high resolution, several H atoms, several deviations from standard geometry and double conformations of several residues are observed in the atomic resolution structure.

The specificity of IDD552 can be associated with the binding of the inhibitor into the specificity pocket (between residues Trp111 and Leu300). The hydrophobic interactions observed in this region are similar to those observed in other ALR2–inhibitor complexes, in particular ALR2–DD594 (Howard *et al.*, 2004).

The main difference from previous ALR2–inhibitor complexes is the pH value during the crystallization: this is the first complex to be crystallized at pH 8 (previous complexes were crystallized at pH 5), thus opening a new area in the crystallographic studies of this enzyme.

A result of importance is the observation of a double conformation of the carboxylic acid moiety of the inhibitor.

In both conformations there is a hydrogen bond between the carboxylic moiety of the inhibitor and the hydroxyl group of Tyr48, but the characteristics of these two bonds differ. The conformation with higher occupancy factor is located next to Tyr48 and is better placed for direct proton donation from this last residue, but does not show any peak in $(F_o - F_c)$ electron-density maps. Furthermore, a water molecule was found for the first time in the ALR2 active site, establishing a water bridge between this inhibitor configuration and Trp111. The molecular-dynamics study shows that the water molecule keeps its position over 750 ps, without moving into or out of the binding site. It suggests the conformations do not interchange over this time period.

In the lower occupancy conformation the geometry is similar to that observed in previous studies (*e.g.* the ALR2–IDD594 complex; Howard *et al.*, 2004), but the peak corresponding to the H-atom position appears to be approximately in the middle of the hydrogen bond. In previously reported structures of AR–inhibitor complexes at subatomic resolution, the hydrogen-peak position appears at the standard position for a protonated tyrosine. It has been argued in previous studies, using a configuration similar to the lower occupancy conformation presented here, that Tyr48 is the proton donor, with His110 acting as a proton relay (Cachau *et al.*, 2000). In this structure, the central position of the H-atom peak suggests that the H atom is shared between the heteroatoms, as observed in other atomic structures (Kuhn *et al.*, 1998) and therefore also opens the possibility of direct proton donation from Tyr48. This observation could be linked to the crystallization pH of 8, a value closer to the pK_a value of Tyr48, and points out the importance of this AR–inhibitor complex structure determination under different crystallization conditions.

The possibility of two inhibitor configurations (extrapolated to the substrate) opens the way for new hypotheses in the ALR2 mechanism, with the possibility of substrate movements into the active site.

We thank the staff of the SBC at APS and the staff (in particular Jean-Luc Ferrer) of the BM30, ESRF, Grenoble for their help in data collection, the IGBMC staff for support with purification, crystallization and computing facilities, Bernard Chevrier for the refinement at 1.8 Å resolution, Roland Stote (ISIS) for his help during the pK_a calculation and Tania Petrova for fruitful discussions. We also thank the reviewers for their many useful suggestions. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the collaboration CERC–CNRS, Ecos Sud, the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Hôpital Universitaire de Strasbourg (HUS).

References

Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. & Karplus, M. (1983). *J. Comput. Chem.* **4**, 187–217.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cachau, R., Howard, E., Barth, P., Mitschler, A., Chevrier, B., Lamour, V., Joachimiak, A., Sanishvili, R., Van Zandt, M., Sibley, E., Moras, D. & Podjarny, A. (2000). *J. Phys. IV Fr.* **10**, Pr10-3–Pr10-13.

Calderone, V., Chevrier, B., Van Zandt, M., Lamour, V., Howard, E., Poterszman, A., Barth, P., Mitschler, A., Lu, J., Dvornik, D., Klebe, G., Kraemer, O., Moorman, A., Moras, D. & Podjarny, A. (2000). *Acta Cryst.* D**56**, 536–540.

Costantino, L., Rastelli, G., Gamberini, M. & Barlocco, D. (2000). *Exp. Opin. Ther. Patents*, **10**, 1245–1262.

Dauter, Z., Lamzin, V. & Wilson, K. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.

Davis, M., Madura, J., Luty, B. & McCammon, J. A. (1991). *Comput. Phys. Commun.* **62**, 187–197.

Deacon, A., Gleichmann, T., Kalb, A., Price, H., Raftery, J., Bradbrook, G., Yariv, J. & Helliwell, J. (1997). *J. Chem. Soc. Faraday Trans.* **99**, 4305.

DeLano, W. (2002). *PyMol.* DeLano Scientific, San Carlos, CA, USA.

El-Kabbani, O., Darmanin, C., Schneider, T., Hazemann, I., Ruiz, F., Oka, M., Joachimiak, A., Schulze-Briese, C., Tomizaki, T., Mitschler, A. & Podjarny, A. (2004). In the press.

Fogolari, F., Brigo, A. & Molinari, H. (2002). *J. Mol. Recognit.* **15**, 377–392.

Frisch, M., Trucks, G., Schegel, H., Scuseria, G., Robb, M., Cheeseman, J., Zakrzewski, V., Montgomery, J., Stratmann, R. & Burant, J. (1998). *Gaussian*98. Gaussian Inc., Pittsburgh, PA, USA.

Gibas, C. & Subramaniam, S. (1996). *Biophys. J.* **71**, 138–147.

Harrison, D., Bohren, K., Petsko, G., Ringe, D. & Gabbay, K. (1997). *Biochemistry*, **36**, 16134–16140.

Howard, E., Sanishvili, R., Cachau, R., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T., Joachimiak, A. & Podjarny, A. (2004). In the press.
Jez, J., Bennett, M., Schlegel, B., Lewis, M. & Penning, T. (1997). *Biochem. J.* **326**, 625–636.
Jez, J. & Penning, T. (2001). *Chem.-Biol. Interact.* **130**, 499–525.
Jorgensen, W., Chandrasekar, J., Madura, J., Irnpey, R. & Klein, M. (1983). *J. Chem. Phys.* **79**, 926–935.
Kinoshita, J. & Nishimura, C. (1988). *Diabetes Metab. Rev.* **4**, 323–337.
Kleywegt, G. & Jones, M. (1996). *Structure*, **4**, 1395–1400.
Kuhn, P., Knapp, M., Soltis, S., Ganshaw, G., Thoene, M. & Bott, R. (1998). *Biochemistry*, **37**, 13446–13452.
Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorsselaer, A., Podjarny, A. & Moras, D. (1999). *Acta Cryst.* D**55**, 721–723.
Leach, A. (2001). *Molecular Modelling. Principles and Applications.* Harlow: Pearson Education.
Longhi, S., Czjzek, M. & Cambillau, C. (1998). *Curr. Opin. Struct. Biol.* **8**, 730–737.
McKerell, A. *et al.* (1998). *J. Comput. Chem. B*, **102**, 3586–3616.
McRee, D. (1999). *J. Struct. Biol.* **125**, 156–165.
Malamas, M. & Hohman, T. (1994). *J. Med. Chem.* **37**, 2059–2070.
Minasov, G., Wang, X. & Shoichet, B. (2002). *J. Am. Chem. Soc.* **124**, 5333–5340.
Miyamoto, S. (2002). *Exp. Opin. Ther. Patents*, **12**, 621–631.
Morris, A., MacArthur, M., Hutchinson, E. & Thornton, J. (1992). *Proteins*, **12**, 345–364.
Nielsen, J. E. & McCammon, J. A. (2003). *Protein Sci.* **12**, 313–326.
Oates, P. & Mylari, B. (1999). *Exp. Opin. Invest. Drugs*, **8**, 1–25.
Oka, M., Matsumoto, Y., Sugiyama, S., Tsuruta, N. & Matsushima, M. (2000). *J. Med. Chem.* **43**, 2479–2483.
Ondrechen, M. J., Clifton, J. G. & Ringe, D. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 12473–12478.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pavelites, J., Bash, P., Gao, J. & McKerell, A. (1997). *J. Comput. Chem.* **18**, 221–239.
Pfeifer, M., Schumer, M. & Gelber, D. (1997). *Diabetes*, **46**, S82–S89.
Ridder, I., Rozeboom, H. & Dijkstra, W. (1999). *Acta Cryst.* D**55**, 1273–1290.
Schlegel, B., Jez, J. & Penning, T. (1998). *Biochemistry*, **37**, 3538–3548.
Schmidt, A. & Lamzin, V. (2002). *Curr. Opin. Struct. Biol.* **12**, 698–703.
Sheldrick, G. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. & Schneider, T. (1997). *Methods Enzymol.* **277**, 319–343.
Simonson, T., Archontis, G. & Karplus, M. (1997). *J. Phys. Chem. B*, **101**, 8349–8362.
Urzhumtsev, A., Tête-Favier, F., Mitschler, A., Barbanton, J., Barth, P., Urzhumtseva, L., Biellmann, J.-F., Podjarny, A. & Moras, D. (1997). *Structure*, **5**, 601–612.
Varnai, P. & Warshel, A. (2000). *J. Am. Chem. Soc.* **122**,3849–3860.
Vlijmen, H. W. T. van, Curry, S., Schaefer, M. & Karplus, M. (1998). *J. Mol. Biol.* **275**, 295–308.
Vlijmen, H. W. T. van, Schaefer, M. & Karplus, M. (1998). *Proteins Struct. Funct. Genet.* **35**, 154–158.
Warren, J., Murdock, G., Ma, Y., Goodman, S. & Zimmer, W. (1993). *Biochemistry*, **32**, 1401–1406.
Wilson, D., Bohren, K., Gabbay, K. & Quiocho, F. (1992). *Science*, **257**, 81–84.
Wilson, D., Tarle, I., Petrash, J. & Quiocho, F. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9847–9851.
Zeng, J., Treutlein, H. & Simonson, T. (1999). *Proteins Struct. Funct. Genet.* **35**, 89–100.