A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil

A Urzhumtsev[1†], F Tête-Favier[1‡], A Mitschler[1], J Barbanton[2], P Barth[3], L Urzhumtseva[1†], J-F Biellmann[3], AD Podjarny[1] and D Moras[1*]

Background: Aldose reductase (AR) is an NADPH-dependent enzyme implicated in long-term diabetic complications. Buried at the bottom of a deep hydrophobic cleft, the NADPH coenzyme is surrounded by the conserved hydrophilic residues of the AR active site. The existence of an anionic binding site near the $NADP^+$ has been determined from the structures of the complexes of AR with citrate, cacodylate and glucose-6-phosphate. The inhibitor zopolrestat binds to this anionic site, and in the hydrophobic cleft, after a change of conformation which opens a 'specificity' pocket.

Results: The crystal structures of the porcine AR holoenzyme and its complexes with the inhibitors tolrestat and sorbinil have been solved; these structures are important as tolrestat and sorbinil are, pharmaceutically, the most well-studied AR inhibitors. The active site of the holoenzyme was analyzed, and binding of the inhibitors was found to involve two contact zones in the active site: first, a recognition region for hydrogen-bond acceptors near the coenzyme, with three centers, including the anionic site; and second, a hydrophobic contact zone in the active-site cleft, which in the case of tolrestat includes the specificity pocket. The conformational change leading to the opening of the specificity pocket upon tolrestat binding is different to the one seen upon zopolrestat binding; this pocket binds inhibitors that are more effective against AR than against aldehyde reductase.

Conclusions: The active site of AR adapts itself to bind tightly to different inhibitors; this happens both upon binding to the inhibitor's hydrophilic heads, and at the hydrophobic and specificity pockets of AR, which can change their shape through different conformational changes of the same residues. This flexibility could explain the large variety of possible substrates of AR.

Addresses: [1]UPR de Biologie Structurale 9004, IGMBC, CNRS/INSERM/ULP, 1 rue Laurent Fries, B.P. 163, 67404 Illkirch, France, [2]LIPHA Industries, 115 Av. Lacassagne, Lyon, France and [3]Laboratoire de Chimie Organique Biologique, Institut de Chimie, Université Louis Pasteur, 4 rue Blaise Pascal, 67008 Strasbourg Cedex, France.

[†]Permanent address: IMPB RAN, Pushchino, Moscow Region, 142292, Russia.

[‡]Present address: Laboratoire de Cristallographie et Modelisation des Materiaux Mineraux et Biologiques, URA CNRS 809, B.P. 239, 54506 Vandoeuvre les Nancy Cedex, France.

*Corresponding author.
E-mail: moras@igbmc.u-strasbg.fr

Key words: aldose reductase, crystal structure, diabetes, inhibitors

Received: 1 January 1997
Revisions requested: 4 February 1997
Revisions received: 28 February 1997
Accepted: 17 March 1997

Structure 15 May 1997, 5:601–612
http://biomednet.com/elecref/0969212600500601

Introduction

Although the physiological role of aldose reductase (AR; EC1.1.1.21) in healthy individuals is still a matter of debate, the enzyme is believed to be of primary importance in the development of severe degenerative complications of diabetes mellitus, through its ability to reduce excess D-glucose into D-sorbitol in noninsulin-dependent tissues [1]. AR has been shown to act on a wide range of substrates *in vitro*, such as aldehydes, aldoses and corticosteroids, but it is most effective on steroid hormones [2], as confirmed by a recent report [3]. This NADPH-dependent enzyme reduces a carbonyl oxygen to a hydroxyl in an ordered 'bi-bi' mechanism, in which NADPH is bound first and $NADP^+$ released last. Like all other NAD(P)-dependent enzymatic reactions described to date, this reduction is stereospecific with respect to the coenzyme. Upon formation of the enzyme–coenzyme–substrate ternary complex, a hydride is transferred from the C4 carbon atom of the nicotinamide ring (hydrogen 4-pro-R of the A-face of the ring) of NADPH [4,5] to the carbonyl carbon of the substrate while a proton is provided by the enzyme to the carbonyl oxygen.

Several crystallographic analyses have been performed on AR; these include the solution of the structure of the apoenzyme from pig lens [6] and of the human [7,8] and pig [9] holoenzymes. These structural studies have revealed that AR folds as a $(\beta/\alpha)_8$ barrel, and have also shown that its catalytic site lies at the bottom of a deep hydrophobic cleft.

Inhibiting AR would provide a way of avoiding the complications of diabetes, and identifying inhibitors is, therefore, an important pharmaceutical goal. Various structurally diverse compounds have been observed to inhibit this enzyme, and they can be divided into four main classes

according to their structures. The first class comprises flavonoids such as quercetin [10]. The second class are spirohydantoins, the most well known of which are sorbinil [11] and alconil [12]. The third class are substituted acetic acids: different subclasses of these are exemplified by 1,3-dioxo-1H-benz[de]isoquinolines, such as alrestatin [13], which was the first aldose reductase inhibitor found to be orally active, by naphtalene derivatives such as tolrestat [14], which is intensively studied, including in large scale trials [15], by 3,4-dihydro-4-oxophtalazines such as ponalrestat [16] and zopolrestat [17], and by rhodanines such as epalrestat [18]. The fourth class of AR inhibitors includes phenylsulfonylnitromethane derivatives such as ICI 215918 [19] and ZD5522 [20].

In this context, the structures of the complexes of AR with the potent inhibitor zopolrestat [21] and with citrate, cacodylate and glucose 6-phosphate have been determined [22]. These analyses have shown that all these molecules bind in the AR active site, and that they recognize a specific anionic site delineated by the C4N of the nicotinamide coenzyme, the Oη of Tyr48 and the Nε of His110. The inhibitor zopolrestat binds in a 'specificity' pocket in the active site, which is accessible after a conformational change of the enzyme.

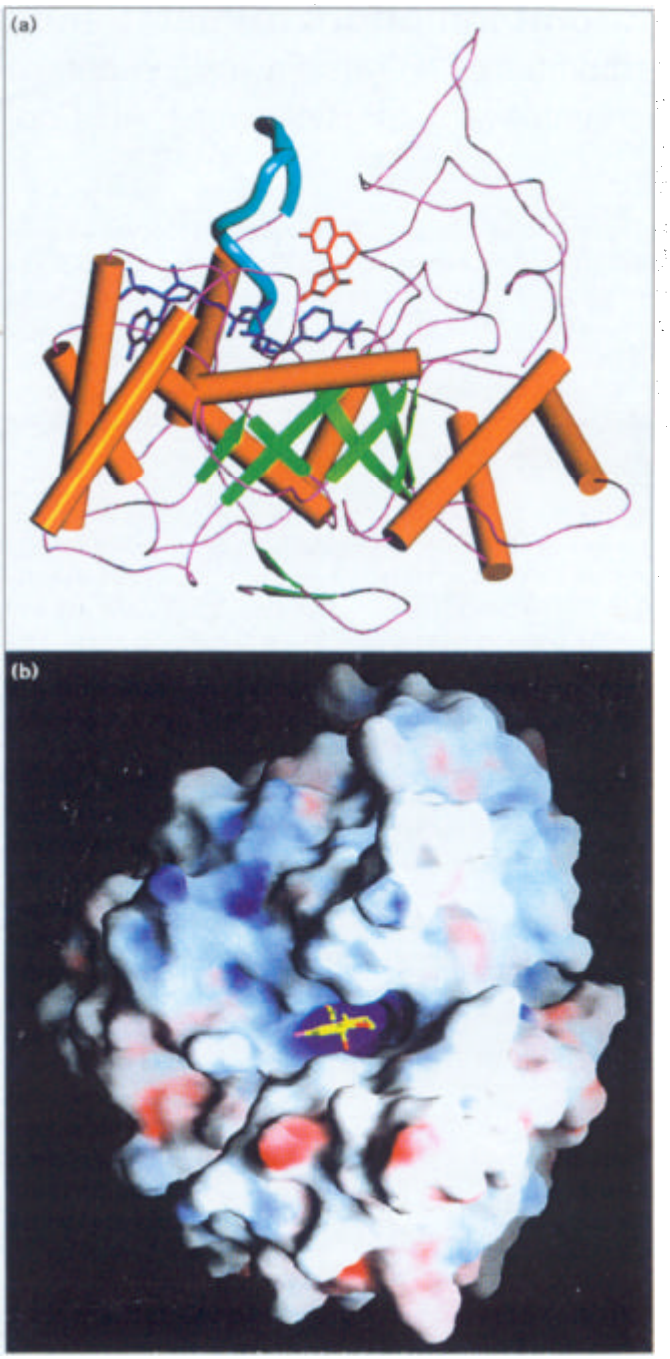
In this work, we report the solutions of the crystal structures of complexes of the pig AR holoenzyme with the noncompetitive inhibitors tolrestat and sorbinil. The present study shows that both of these inhibitors bind in the active site; the contacts between the inhibitors and the active-site residues are analyzed. A kinetic explanation for the apparent contradiction between the active site binding and non-competitiveness of AR inhibitors has been given by Nakano and Petrash [23] for the case of zopolrestat. We find that the inhibitors tolrestat and sorbinil recognize the same anionic site as zopolrestat, and that tolrestat binds in the same specificity pocket as zopolrestat. The orientations of the hydrophobic moieties of tolrestat and zopolrestat in this pocket are, however, quite different, and the conformation of the specificity pocket varies according to the inhibitor that is bound.

Results and discussion

The enzyme structure

The structure of the pig AR holoenzyme [9] has now been obtained at 2.8 Å resolution from a tetragonal crystal form with one monomer per asymmetric unit, and further refined to 2.0 Å resolution (see Materials and methods). This structure, which is very similar to that of the human holoenzyme, folds as a $(\beta/\alpha)_8$ barrel, with the coenzyme binding on the top of the barrel (Fig. 1a) and the active site at the bottom of a deep hydrophobic pocket. The very high structural similarity of the pig and human enzymes might possibly extend to the other ARs of known sequence (bovine lens, rabbit kidney, rat lens and mouse

Figure 1



Overall views of the holoenzyme form of pig lens AR. **(a)** Schematic drawing obtained using the program SETOR [37], showing the TIM barrel structure (green), the sorbinil molecule (brown ball-and-stick) and the $NADP^+$ (dark blue ball-and-stick) lying between strands S7 and S8, with loop L7 in light blue bent over it. **(b)** View of the surface of the molecule generated using the program GRASP [38], with the sorbinil molecule (ball-and-stick) bound in the active-site cleft. The colors correspond to electrostatic potential surfaces (positive in blue; negative in red) calculated using charge distributions on all protein residues using the program XPLOR [32], and on $NADP^+$, from Pavelites *et al.* [39]. Note the positively charged bottom of the pocket, below the sorbinil molecule.

Table1

Principal interactions of the residues in the hydrophilic network around nicotinamide (Fig. 2a).

Residue	Active-site contacts			Sorbinil	Sorbinil	Tolrestat	Tolrestat
	H-bond	VdW	Other	VdW	H-bond	VdW	H-bond
Asp43	Gln183 NADP+		Salt bridge Lys77	–	–	–	–
Tyr48	Lys77 Wat1			6	2	4	1
Lys77	Tyr48 Cys44 Wat2	His110	Salt bridge Asp43	–	–	–	–
His110	Asn160 Wat1 Wat2	Trp79 Lys77 Trp111		6	1	6	2
Ser159	NADP+ Asn160			–	–	–	–
Asn160	NADP+ Ser159 His110 Wat3	Trp111		–	–	–	–
Gln183	Asp43 NADP+			–	–	–	–
Tyr209	Wat3		Charge-transfer complex NADP+	–	–	–	–

The type of interaction – hydrogen bond (H-bond), van der Waal (VdW), or other – is indicated, as well as the number of inhibitor–protein contacts.

kidney), as they each have more than 84% sequence identity with the pig enzyme.

As NADP+ was present during crystallization, which was done under mildly acidic conditions, it is likely that the form of AR crystallized is the complex with NADP+. Further evidence for this idea comes from the similarity of the structure of the pig holoenzyme to that of the known structure of the human holoenzyme, which is AR complexed with NADP+, and from the fact that endogenously bound NADP+ was found in the first tetragonal crystals formed, even though no coenzyme was added during the original purification and crystallization (data not shown).

Coenzyme binding

Comparative analysis of the crystal structures of the apoenzyme [6] and the holoenzyme shows that the binding of AR to NADPH, involving numerous interactions similar to those reported for the human holoenzyme [7], requires significant conformational changes in AR. The most important conformational change concerns loop 7, which spreads over the coenzyme (Fig. 1a); Trp20 and Lys21 within this loop each occupy two different positions in the two forms of AR, corresponding to the opening and closing of a latch formed by the bulky aromatic tryptophan residue and a salt bridge between Lys21 and Asp216. These conformational changes probably account for the ordered bi-bi mechanism displayed by the enzyme and also possibly for its low turnover. Indeed, it has been shown that NADP+ release is the rate-limiting step of the AR-catalyzed reduction of D-glyceraldehyde [24,25].

The active site

The nicotinamide ring of the coenzyme is largely buried at the bottom of a deep cleft, which is approximately 4×15 Å wide and 15 Å deep (Fig. 1b), situated at the C-terminal side of the cylindrical sheet above strands S1 and S2. The only accessible atoms of NADPH are C5, and the hydride donor C4, thus defining the cleft as the only possible active site. The residues lining the cleft can be divided into two groups (Tables 1 and 2).

The first group comprises the, mainly hydrophilic, residues Asp43, Tyr48, Lys77, His110, Ser159, Asn160, Gln183 and Tyr209 (Table 1). Totally conserved among ARs, these residues form a tight network of hydrogen bonds linked to the nicotinamide coenzyme (Fig. 2a). Ser159, Asn160 and Gln183, which are hydrogen-bonded to the amide group of the nicotinamide, assure the precise positioning of the nicotinamide ring for stereoselective transfer of the 4-pro-R hydrogen. This network of hydrogen bonds involves well-ordered water molecules, marked in Figure 2a as Wat1, Wat2 and Wat3. The water molecule Wat1 occupies the anionic site, between Tyr48 and His110. Also noteworthy is the stacking interaction with Tyr209 (Fig. 2a), which suggests stabilization of the positively charged aromatic system formed upon oxidation of NADPH to NADP+

Table 2

Residues at the hydrophobic sides of the active site (A) and the specificity pocket (S).

Residue	Position	Sorbinil VdW	Sorbinil H-bond	Tolrestat VdW	Tolrestat H-bond	Sequence Conservation
Trp20	A	28	1 (N–F)	13	1 (N–S)	Conserved in ARs and AR1s
Trp79	A	2	–	2	–	Conserved in ARs and AR1s
Trp111	A+S	4	1	12	1	Conserved in ARs and AR1s
Phe122	A+S	4	–	6	–	Conserved in ARs and AR1s
Pro218	A	–	–	–	–	Conserved in ARs; Ala in AR1s
Trp219	A	2	–	2	–	Conserved in ARs
Cys298	A	1	–	–	–	Conserved in ARs; Ile in AR1s
Leu300	A+S	2	–	4	1 (O–F)	Conserved in ARs; Pro in AR1s
Val47	A	3	1 (O–F)	1	–	Conserved in ARs; Ile in human AR1 and pig AR1
Thr113	S	–	–	1	–	Conserved in ARs; Tyr in AR1s
Phe115	S	–	–	3	–	Conserved in ARs and AR1s
Val130	S	–	–	2	–	Conserved in ARs; Ile in human AR1 and pig AR1
Ser302	S	–	–	4	1 (O–F)	Conserved in ARs; extra loop in AR1s
Cys303	S	–	–	2	3 (N–F, S–F, S–O)	Conserved in ARs; extra loop in AR1s

AR–inhibitor interactions — number of hydrogen bonds (H-bond) or van der Waals (VdW) contacts — and sequence conservation in mammalian ARs and aldehyde reductases (AR1s) are shown.

(observed in the holoenzyme crystal form) by a charge-transfer complex.

The second group of residues lining the cleft (Table 2), all of which lie above the coenzyme, includes those that form the walls of the hydrophobic cleft — Trp20, Trp79, Trp111, Phe122, Pro218, Trp219, Cys298, Leu300 and Val47. Trp20, Phe122 and Trp219 fully face the active site, and are thus expected to make the major contacts with a potential substrate, as confirmed by the structure of the complexes of AR with tolrestat and sorbinil described below. The structure of the AR–zopolrestat complex also shows the existence of a specificity pocket, the residues of which are also in the second group of cleft-lining residues and include Thr113, Phe115, Val130, Ser302 and Cys303. Residues Trp111, Phe122 and Leu300 are shared by both the active site and the specificity pockets. These pockets are occupied by a network of water molecules, marked Wat1 to Wat11 in Figure 2b.

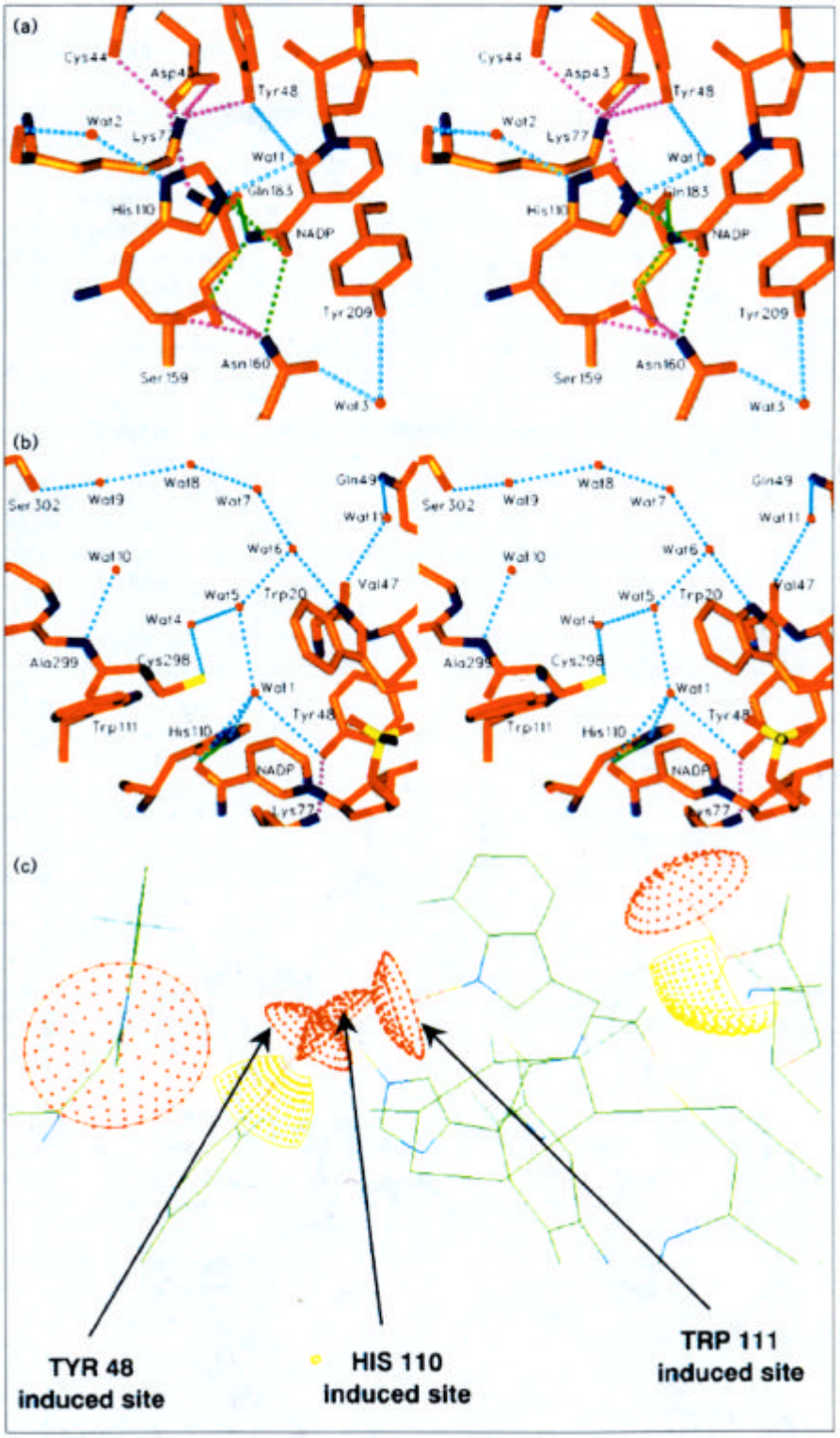
Modeling calculations performed by De Winter and von Itzstein [26] showed that the protein residues involved in possible hydrogen-bonding interactions with a substrate at the bottom of the AR active site are Tyr48, His110 and Trp111. Following their work, we have performed a theoretical calculation of possible binding sites using the program FLEXX [27]. It showed that the spherical caps corresponding to the positions of possible hydrogen-bond acceptors of these three residues intersect (Fig. 2c), joining in an extended surface centered on the anionic site between His110 and Tyr48, where the intersection is largest. This surface, with a neighboring volume allowing for differences in hydrogen-bond length, can be considered to be the recognition region for hydrogen-bond acceptors.

Inhibitor binding

Complexes of the AR holoenzyme and the inhibitors tolrestat and sorbinil were prepared by soaking the holoenzyme crystals in a solution of inhibitor (see Materials and methods). Following the studies by Nakano and Petrash [23], binding of both inhibitors to the $NADP^+$ and NADPH forms of the holoenzyme was confirmed by mass spectroscopy (A van Dorsselaer, personal communication).

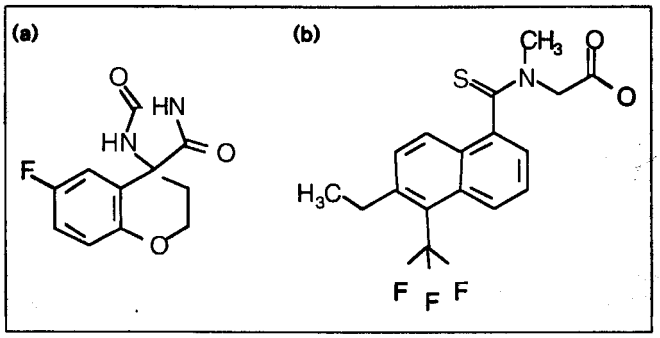
Figure 2

Details of the active site of pig lens AR. **(a)** The group 1 residues and $NADP^+$ are shown in ball-and-stick form; the network of hydrogen bonds is shown in dotted lines. The water molecule Wat1 occupies the anionic center. Note the stacking of Tyr209 and the nicotinamide ring; the arrangement of the rings corresponds to the geometry of a charge-transfer complex. **(b)** View perpendicular to (a), showing the network of water molecules Wat1–Wat11 occupying the active-site cleft. **(c)** View of the bottom of the active site, generated using the program FLEXX [27], showing the possible sites for binding hydrogen-bond acceptors (red caps) and hydrogen-bond donors (yellow caps). The protonation state of His110 at Nε2 is the one proposed by De Winter and von Itzstein [26]; our current work with inhibitor docking tends to confirm this choice (unpublished observations).



These inhibitors have a polar 'head' and a hydrophobic ring system (Fig. 3). We generated difference maps, and these clearly showed that the inhibitors bind into the active site, with the polar head near the coenzyme (Fig. 4). Inhibitor models were positioned in these difference maps and the complexes were refined (see Materials and

Figure 3



Two dimensional chemical representation of **(a)** sorbinil and **(b)** tolrestat.

methods); in the case of sorbinil, there were no significant changes in the enzyme structure, and in the case of tolrestat, there was a displacement of Leu300 and Phe122. Tables 1 and 2 show the contacts of the inhibitors with the enzyme, which can be classified into hydrophilic contacts at the bottom of the site, hydrophobic contacts and water-mediated contacts.

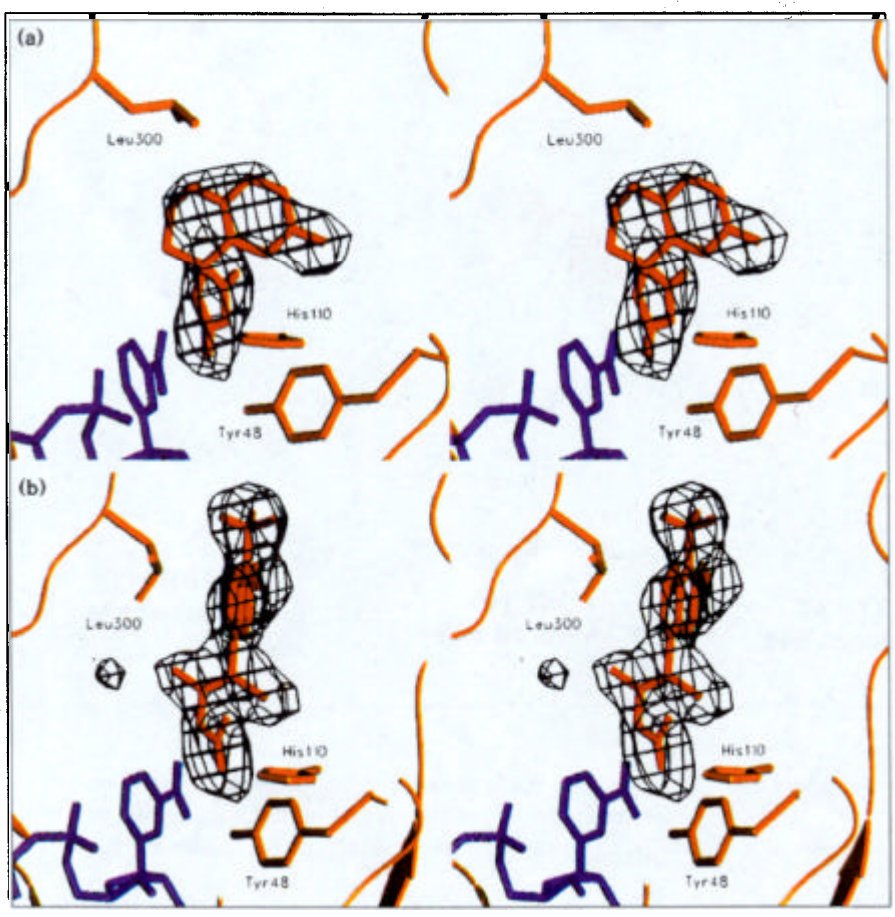
Hydrophilic contacts at the bottom of the site

Sorbinil

In the case of sorbinil, there are two carboxyl oxygens, O2 and O3, with an NH group in the middle; the oxygens and nitrogen make a tight hydrogen-bonding network with Tyr48, His110 and Trp111 of AR (Fig. 5a), and are in close contact with the bottom of the site: firstly, the O2 atom is between the OH group of Tyr48 (hydrogen bond), the C4N of nicotinamide and the Cγ of Trp20 (at distances of 2.7 Å, 3.2 Å and 3.1 Å, respectively); secondly, the N2 atom occupies a middle site at 2.9 Å from the Nε2 of His110, 3.2 Å from the OH group of Tyr48 and 3.5 Å from the C4N of nicotinamide; and thirdly, the O3 atom is at 3.0 Å from the Nε2 of His110 (hydrogen bond) and 2.9 Å from the Nε1 of Trp111 (hydrogen bond). The sorbinil could also be in a second tautomeric state, in which the hydrogen is on O2 or O3 and the N2 atom has a lone pair of electrons; in this case, the N2 atom would make a hydrogen bond to His110 and Tyr48.

These three atoms (O2, O3 and N2 of sorbinil) mark three possible hydrophilic binding centers on AR: one between Trp20 and Tyr48, the second between Tyr48 and His110

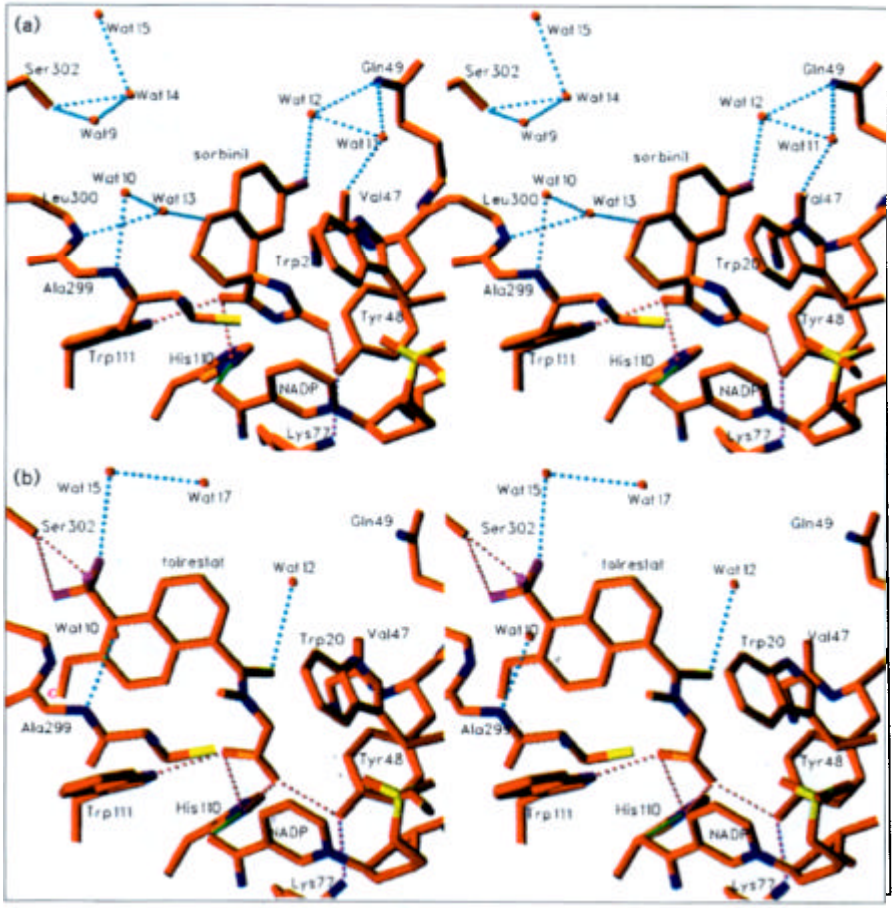
Figure 4



Difference maps showing **(a)** sorbinil and **(b)** tolrestat in the active site of AR. The difference maps F_o–F_c are calculated at 3 Å, using the phases from the native enzyme. For sorbinil, the density indicates clearly that the *S*-enantiomer is bound, in agreement with the inhibition studies in solution [1]. Note the conformational change of the sidechain of Leu300, necessary to accommodate the tolrestat molecule.

Figure 5

Hydrophilic binding site. **(a)** Binding of sorbinil to the hydrophilic active site of AR. Note the tight hydrogen-bond network (dashed lines). Atoms O2, N2 and O3 mark the three possible hydrophilic centers. **(b)** Binding of tolrestat to the hydrophilic active site. Note the tight hydrogen-bond network (dashed lines).



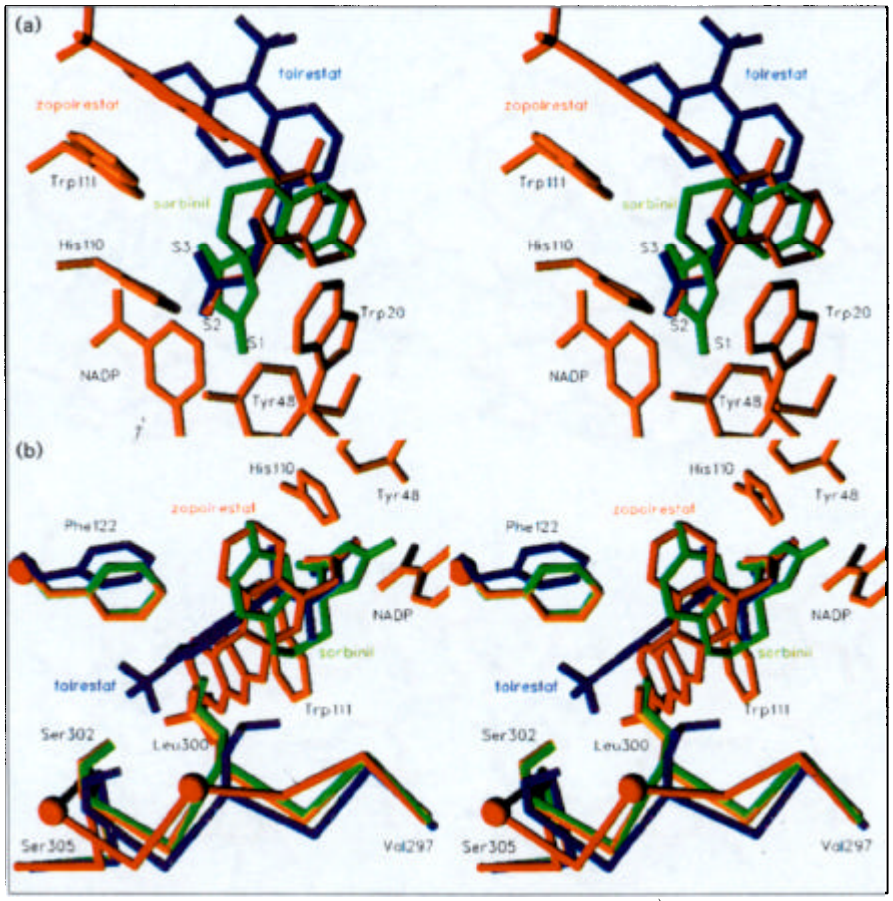
(corresponding to the anionic site) and the third between His110 and Trp111 (Figs 5a and 6a). All three centers are inside the recognition region defined theoretically using the program FLEXX; centers 2 and 3, which can make two hydrogen bonds, are in the region of the intersection of the corresponding spherical caps (see Fig. 2c). A representation of the protein surface with the electrostatic potential shows that all three centers are on top of a positively charged region, which is quite extended and centered below the anionic site (Figs 1b and 7a).

Tolrestat

Tolrestat has a COOH hydrophilic head, similar to that of many AR substrates. Given the p*K*a of tolrestat (4.5), this head is probably bound by AR in the form $(COO)^-$, with the negative charge shared between atoms O2 and O3. As in the case of sorbinil, these atoms are bound to Tyr48, His110 and Trp111 through a tight hydrogen-bond network, and are in close contact with the bottom of the active site (Figs 5b and 6a): firstly, O2 is at a distance of 2.7 Å from the Oη of Tyr48 (hydrogen bond), 2.6 Å from the Nε2 of His110 (hydrogen bond) and 3.2 Å from the C4N of nicotinamide; and secondly, O3 is at 2.8 Å from the Nε1 of Trp111 (hydrogen bond) and 3 Å from the Nε2 of His110 (hydrogen bond). Note that the hydrogen-bond donor character of the Oη of Tyr48 is accentuated by the presence of Lys77; furthermore, the C4N atom of nicotinamide is 3.3 Å from the C16 atom of tolrestat. Referring to the hydrophilic binding centers of AR defined by sorbinil, both centers 2 and 3 are occupied by the negatively charged oxygens O2 and O3 of tolrestat, respectively.

Analysis of the structures of AR complexes with zopolrestat (Fig. 6a), cacodylate and glucose 6-phosphate show that, in these structures, the hydrophilic binding centers 2 and 3 of AR are also occupied. In the case of the AR complex with citrate, centers 1 and 2 are occupied by negatively charged oxygens, and there is an extra interaction with Trp111. Our analysis of eight other inhibitor complexes shows occupation of the same centers, with a preference for center 2 (unpublished observations). Therefore, all three centers can accommodate negative charges depending on the structure of the inhibitor; center 2, which corresponds to the anionic site, is clearly the most occupied one.

Figure 6



Comparison of bound inhibitors. **(a)** Superposition of the inhibitors sorbinil (green), tolrestat (blue) and zopolrestat (red) with the residues of the hydrophilic active site of AR (orange). Note how all three inhibitors share centers S2 and S3, and that sorbinil occupies center S1. **(b)** Superposition of the complexes of AR with sorbinil (green), tolrestat (blue) and zopolrestat (red) with the native protein (orange). Note the changes of the sidechains of Leu300, Phe122 and Ser302 for the tolrestat structure. For zopolrestat, only the Cα atoms of the AR residues are shown; note their displacement.

Hydrophobic contacts and water-mediated interactions

Sorbinil

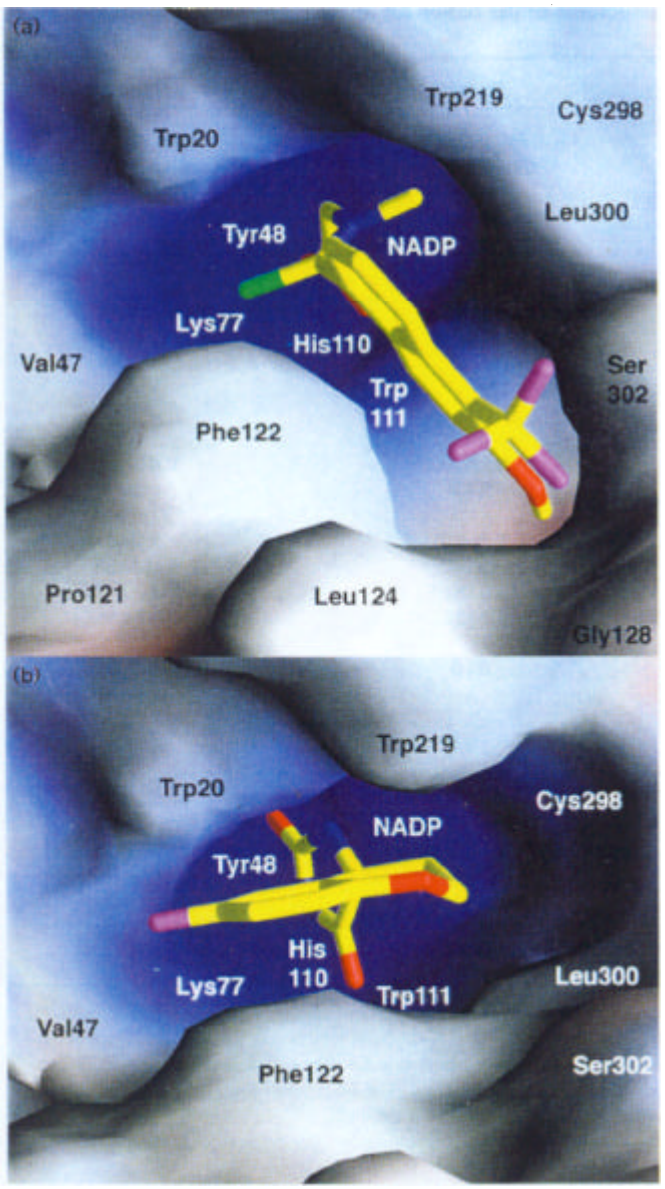
The sorbinil molecule occupies the bottom of the active site, between Trp20 on one side and Trp111–Phe122–Leu300 on the other. The ring system is stacked against Trp20, making 29 contacts (Fig. 6a). The binding of sorbinil displaces the water network formed by Wat1 and Wat4–Wat8 that occupied the active site. On the other hand, two new water molecules, Wat12 and Wat13 (Fig. 5a) link the sorbinil molecule to the protein.

Tolrestat

The chain between C13 and C16 of tolrestat emerges from the anionic sites making mostly van der Waals contacts with Trp20 of AR (Fig. 6a), with a possible hydrogen bond between the S1 atom of the tolrestat molecule and Nε1 of Trp20 (3.8 Å). The hydrophobic ring system of tolrestat gets trapped tightly in a pocket formed by Trp111, Leu300 and Phe122, making 13 contacts with Trp111 (Fig. 6b). The opening of this pocket (the specificity pocket; see below) requires a conformational change at the level of the sidechains of Leu300 and Phe122. The O1 end of the tolrestat molecule is in contact with the C4 atom of Phe115 of AR, while the F atoms pack against Ser302 and Cys303, with an OH–F bond with the OH of Ser302 (Fig. 6b). The binding of tolrestat displaces the water network formed by Wat1 and Wat4–Wat9 that occupied the active site and the specificity pocket.

The hydrophobic pockets marked by these two inhibitors are, therefore, the bottom of the active site (between Trp20 and Trp111–Phe122) and the specificity pocket (between Trp111, Leu300 and Phe122). These two pockets have already been noted in the structure of the AR–zopolrestat complex, where the first one traps the phtalazinone ring and the second one traps the benzothiazole ring, making 65 contacts altogether. Note that the mutation of Trp20 into alanine stops sorbinil inhibition, has a large effect on tolrestat inhibition and has a less marked effect on zopolrestat inhibition [28]. This is consistent with the relative importance of the two hydrophobic pockets in binding the different inhibitors: sorbinil is bound mostly to Trp20 (29 contacts) and does not open the specificity pocket, whereas tolrestat makes an equal number of contacts to Trp20 and Trp111 and zopolrestat is bound mostly to Trp111.

Figure 7



Changes in the AR molecular surface corresponding to inhibitor binding. **(a)** Superposition of tolrestat on a molecular surface of the protein; note that the conformational changes have opened the specificity pocket, which is occupied by the inhibitor. **(b)** Superposition of sorbinil on a molecular surface of the AR protein, in the same orientation as in (a); note that the inhibitor occupies the active-site pocket. Figure was generated using the program GRASP [38].

The specificity pocket

As noted above, the binding of both tolrestat and zopolrestat requires a conformational change in AR that opens the specificity pocket, which is contiguous with the active-site pocket. In the case of the native enzyme or the AR–sorbinil complex, the specificity pocket is closed, (Fig. 7b) and the residues Trp111, Phe122, and Leu300 form one hydrophobic wall of the active site pocket, with Phe122 and Leu300 4Å away (Fig. 6b). In the case of the AR–zopolrestat complex, the loops 121–135, containing Phe122, and 299–302, containing Leu300, move in concert [21], so that the Leu300–Phe122 contact is kept but they shift away from Trp111, creating a space for the zopolrestat molecule (Fig. 6b). In the case of the AR–tolrestat complex, only the sidechains of Leu300 and Phe122 move significantly and apart from each other (Fig. 6b), creating a space for placing the tolrestat molecule at a position perpendicular to that occupied by zopolrestat in the AR–zopolrestat complex (Figs 6b and 7a).

Therefore, even though the specificity pocket involves the same residues in both cases, the way it opens to bind zopolrestat and tolrestat is different, and leads to distinct orientations of the bound inhibitors. Zopolrestat is stacked parallel to Trp111, whereas tolrestat packs between Phe122 and Leu300, perpendicular to Trp111; zopolrestat goes more deeply into the pocket, with one F atom hydrogen-bonded to Thr113, whereas in tolrestat the F atom is hydrogen-bonded to Ser302, much closer to the pocket opening. In short, while they share the same pocket, binding of tolrestat requires much fewer structural changes to AR than binding of zopolrestat.

A role for the specificity pocket

It is interesting to note that while sorbinil inhibits both AR and aldehyde reductase with similar IC_{50} values [29] (5.4 μM for aldehyde reductase and 2 μM for AR), tolrestat and zopolrestat inhibit AR more than than inhibit aldehyde reductase (the IC_{50} values are 0.72 and 27.0 μM for aldehyde reductase and 0.01 and 0.06 μM for AR, respectively) [29]. This strongly suggests that, whereas sorbinil binds both enzymes in a similar way, tolrestat and zopolrestat bind them differently. We performed a sequence comparison of mammalian ARs and aldehyde reductases in order to analyze the conservation of residues involved in inhibitor binding (see Materials and methods). As shown in Table 2, active site residues that are strongly involved in sorbinil binding are conserved in both enzymes. Whereas most of the specificity pocket residues involved in tolrestat binding tend to be conserved in mammalian ARs, however, they vary in aldehyde reductases. In particular, Leu300 is conserved in ARs but is replaced by Pro301 in aldehyde reductase; Thr113 is conserved in AR, but is replaced in aldehyde reductase by Tyr115; and Ser302 is conserved in AR, but in aldehyde reductase there is an extra loop, from Pro301 to Arg312, inserted at this position. The insertion of this extra loop changes the nature of the pocket and could explain the difference in tolrestat binding. Furthermore, in the case of zopolrestat, the replacement of Thr113 of AR by Tyr115 in aldehyde reductase takes up the space that the inhibitor molecule occupies in the zopolrestat–AR complex.

Ongoing crystallographic studies on inhibitors AL1576 (imirestat) and Amino SNM show that, like sorbinil, they

Table 3

Statistics of data collection and refinement for the solutions of the structures of the native AR enzyme and the complexes of AR with the two inhibitors.

Object	Native	Sorbinil	Tolrestat
Space group	$P4_32_12$ (96)	$P4_32_12$ (96)	$P4_32_12$ (96)
Unit cell: a, c (Å)	68.09, 153.35	67.92, 152.80	68.42, 153.56
Diffraction data:			
resolution	2.0–15.0	2.3–15.0	2.3–15.0
No. reflections	24329	13052	15449
completeness (%)	97	79	92
completeness at the	93	45	66
highest resolution (%)	(2.0–2.1)	(2.3–2.4)	(2.3–2.4)
Shell			
R_{sym} (%)*	5.1	4.9	4.5
Rigid body refinement:			
R_{init} (3–8 Å)†	27.6	39.7	38.9
overall shift	0.14 Å, 0°	0.35 Å, 0.1°	0.46 Å, 0.1°
R_{final} (3–8 Å)†	26.3	24.0	25.4
Refinement:			
resolution	2.0–6.0	2.3–6.0	2.3–6.0
N reflections ($F>2\sigma$)	23487	12147	14456
R_{start} (R_{free})†	26.7 (33.9)	23.0 (31.8)	24.9 (32.5)
R_{fin} (R_{free})†	20.1 (28.5)	19.2 (26.3)	20.7 (29.0)
Final model:			
protein residues	315	315	315
coenzyme	1	1	1
inhibitor	–	1	1
water molecules	277	174	150
Geometrical statistics:			
Δ_{bond} (Å)	0.012	0.011	0.011
Δ_{angle} (°)	2.88	2.78	2.80
Δ_{impr} (°)‡	1.19	1.14	1.17
Δ_{dihe} (°)‡	24.8	24.9	25.1
Ramachandran plot§:			
favourable regions (%)	90.0	90.0	89.3
additional regions (%)	10.0	10.0	10.7

In general, $R=\Sigma(F_1-F_{obs})/\Sigma(F_{obs})$, in which F_{obs} = experimental structure factor amplitude. *For R_{sym}, F_1 = symmetry related experimental structure factor amplitude. †For R_{init}, R_{final}, R_{start} and R_{fin}, F_1 = model structure factor amplitude, corresponding to the beginning of the rigid body refinement, end of the rigid body refinement, beginning of the atomic refinement and end of the atomic refinement, respectively. ‡$\Delta_{dihe}=<\alpha-\alpha_o>$, in which α is the dihedral angle, and $\Delta_{impr}=<\beta-\beta_o>$, in which β is an angular value used by X-PLOR to define chirality or planarity. §The Ramachandran plot was calculated with the program PROCHECK [40].

bind the active site of AR without opening the specificity pocket (our unpublished observations). Since AL1576 is specific for aldehyde reductase [27], our observations further confirm the role of the specificity pocket.

Biological implications

Diabetes is a debilitating disease leading to severe complications and a shortened life expectancy. In noninsulin-dependent tissues, insulin therapy does not prevent the occurrence of these complications, which include neuropathy, retinopathy, nephropathy and cataracts. Under conditions of diabetic hyperglycaemia, excess glucose is metabolized by the polyol pathway, which comprises two enzymes: aldose reductase (AR), which reduces excess D-glucose into D-sorbitol, and sorbitol dehydrogenase, which converts sorbitol to fructose. Diabetic complications have been linked to excessive accumulation of sorbitol, leading to an interest in stopping the polyol pathway through the inhibition of AR. This NADPH-dependent enzyme reduces a carbonyl group to a hydroxyl one in an ordered 'bi-bi' mechanism, in which NADPH is bound first and $NADP^+$ released last. The catalytic site of AR is at the bottom of a deep hydrophobic cleft which recognizes a large variety of substrates, ranging from glucose to steroid hormones. To understand this promiscuous binding, the structures of the complexes of AR with two inhibitors, sorbinil and tolrestat, have been determined. This study has pharmaceutical interest, as tolrestat and sorbinil have been widely tested clinically with the aim of preventing complications associated with diabetes.

Both inhibitors bind into the active site of AR, their polar heads being recognized by an extended system of

three hydrophilic centers near the coenzyme. The hydrophobic ring system of sorbinil binds to the native active site, while that of tolrestat binds into a pocket opened by conformational changes in the walls of the active site cleft. These conformational changes are different from those previously observed for the inhibitor zopolrestat, which binds the same pocket. Inhibitor binding into this 'specificity' pocket is correlated with a tighter binding to AR than to aldehyde reductase, a closely related enzyme.

The conservation of key residues in ARs from different species and their differences from the equivalent residues in aldehyde reductase suggest that the specificity pocket binds substrates unique to AR. The different ways in which the specificity pocket opens leads to tight packing of different inhibitors by the same residues; this property might prove to extend to different substrates. Therefore, the study of these AR–inhibitor complexes and the conformational changes induced by the inhibitors explains some of the remarkable adaptability of the active site of AR, as well as its specificity for certain inhibitors.

Materials and methods

Purification and crystallization of the pig AR holoenzyme

Pig eyes collected from freshly slaughtered animals were immediately frozen on dry ice. The lenses were extracted, frozen on dry ice and stored at −20°C until used. The compounds D-xylose, dithioethreitol (DTT) and MES were obtained from Sigma or Boehringer-Mannheim, NADP and NADPH from Boehringer-Mannheim, xylitol from AR2ich Chemicals, and polyethylene glycol (PEG) 3000 and 6000 and EDTA from Fluka.

The purification [30] was performed at 4°C and all the buffers were degassed with argon. Pig lenses (100 g) were homogenized in a Waring blender for 2 min in 200 ml buffer A (50 mM phosphate buffer pH 7.0, 1 mM EDTA, 5 mM DTT, 0.002% hibitane). The homogenate was centrifuged at 10 000 rpm for 90 min. The supernatant was brought to 30% (w/v) PEG 3000 by adding 1.0 vol of a 60% (w/v) PEG 3000 solution in buffer A, and stirred for 1 h under argon. After centrifugation as above, the supernatant was diluted with water (600 ml) and loaded on a DEAE 50 Sephadex column (200 ml), equilibrated with buffer B (20 mM phosphate buffer pH 7, 1 mM EDTA, 1 mM DTT). The elution was performed first with buffer B (500 ml) and then with a gradient of buffer B (600 ml) and of 0.5 M phosphate buffer pH 7.0, 1 mM EDTA, 1 mM DTT (600 ml). The flow rate was 20 ml h^{-1} and fractions of 10 ml were collected. The fractions with the highest activity were concentrated with an Amicon PM 10 membrane. The concentrated fractions were dialyzed against 25 mM MES buffer pH 6.2, 1 mM EDTA, 5 mM DTT (three times 1l), concentrated in Centricon PM 10 microconcentrators (Amicon), and the activity of the purified enzyme was tested and protein determination was performed.

The crystals of holoenzyme were grown using the hanging-drop method by vapor diffusion with PEG 6000. The drop concentration was 15 mg ml^{-1} AR, 2 equivalents of $NADP^+$ in 75 mM MES buffer pH 6.2, 1 mM EDTA, 1 mM DTT, 2.5% PEG 6000 (w/v), and the well contained the same buffer with 25% PEG 6000 added. Suitable crystals were obtained within 10 days. The complexes of AR with the inhibitors were obtained in the following way: the crystals were transferred to a capillary, the mother liquor was removed and a solution containing the inhibitor was added (120 mM TES buffer pH 6.2, 1 mM EDTA, 1 mM DTT, 30% PEG 6000 (w/v), and either 10 mM sorbinil or 62 mM tolrestat).

Determination of the AR holoenzyme structure at 2.8 Å resolution

The structure of the porcine AR holoenzyme was obtained as follows: tetragonal crystals of AR were obtained at 4°C in MES buffer (25 mM, pH 6.2), using PEG 6000 (20%) as the precipitant. The space group was $P4_32_12$ with unit cell dimensions a = b = 68.5 Å, c = 153.9 Å, and one monomer per asymmetric unit. Data were collected to 2.8 Å resolution from one crystal mounted on a high brilliance rotating anode, using a Siemens area detector, and treated with the XDS package [31]. The overall R_{sym} was 5.6% on intensities for 9093 independent reflections above 3σ, obtained from 46 450 observations (overall completeness above 3σ = 94%).

A mercury derivative was prepared by soaking the crystal in ethylmercury chloride (0.9 mM, 10 days). It revealed three mercury sites, which were found to have the same relative positions and occupancies as the sites of the mercury derivative obtained for the apoenzyme (space group P1). This phenomenon was used to calculate the rotations and translations that place the model of the apoenzyme in the tetragonal cell of the holoenzyme crystals. A few cycles of rigid-body refinement at very low resolution (∞–20 Å) were necessary to slightly correct the position of the model (R factor of 35.5%). Unexpectedly (since no NADPH had been added to the crystallization buffer), the Fourier difference map computed after a few cycles of energy minimization showed clear electron density for endogenously bound $NADP^+$. The complex including the coenzyme was then refined by energy minimization and slow-cooling dynamics (X-PLOR [32]), with a final crystallographic R factor of 21.3% for all the measured reflections in the 8.0–2.8 Å resolution range and of 20.5% for the reflections with amplitudes $F > 2\sigma(F)$ in the same resolution range (97.2% of the total number). No water molecules were included.

Refinement of the AR holoenzyme structure at 2.0 Å resolution and structure determination of the enzyme–inhibitor complexes

Complexes of the holoenzyme with tolrestat or sorbinil were obtained by soaking improved holoenzyme crystals in a solution containing inhibitor. These inhibitor-soaked crystals diffracted at a synchrotron to a resolution of 2 Å and were isomorphous to the improved native ones (see Table 3 for statistics). Diffraction data for the native structure were then measured again using a larger crystal and a synchrotron source. This led to a clear improvement of the limits of the diffraction data, from 2.8 Å to 2.0 Å (see Table 3 for statistics). The previously determined model was used as the starting point of a refinement, which included a rigid-body refinement step, a slow-cooling step, a Powell minimization and a temperature-factor refinement. Water molecules were placed in a difference map.

On the basis of this refined native structure, we performed a series of steps for both AR–inhibitor complexes: first, placement and rigid-body refinement of the native model in the complex unit cell at 3 Å resolution; second, analysis of the difference density maps at 3 Å resolution; third, manual model correction for tolrestat (following difference map); fourth, refinement of the model without an inhibitor at the resolution of 2.3 Å; fifth, insertion of the inhibitor model into the improved density map; and sixth, refinement of the model of the complex including manual corrections. At the step of rigid-body refinement and calculation of the first maps, residues 116–133 and 213–223, which form flexible loops, were removed from the starting model. Their position was observed in the maps to be very close to those in the native model and therefore they could be easily inserted back at the second step.

For both complexes, the density maps showed clearly the presence of the inhibitor in the active center (Fig. 4). The quality of the density was very high and the inhibitor models were built without any difficulty. The programs O [33] and X-PLOR [32] were used for the model building and refinement. Large differences between the protein models which demanded a manual correction were observed for the two N-terminal residues (different for all three structures) and for Leu300 in the tolrestat complex.

The analysis of the weighted density maps, where the phase error estimates from the refined models were calculated by RFLEXPL [34], clearly showed the presence of a number of water molecules. The first sets of water molecules (about 100) were placed automatically and independently for both inhibitors using the program MAPMAN [35]. It is interesting to note that about 90% of these water molecules were placed in similar positions for tolrestat and sorbinil. These water molecule sets were edited, extended and refined together with the models of the complexes.

Accession numbers

The coordinates of the sorbinil complex (accession number 1ah0), the tolrestat complex (accession number 1ah3) and the native holoenzyme (accession number 1ah4) have been deposited in the Protein Data Bank.

Sequence alignment

The known sequences of mammalian ARs (pig, bovine, rat, mouse, rabbit and human) and aldehyde reductases (human, rat and pig) were retrieved from the SwissProt Data Bank. The sequences were aligned using the program PILEUP of the GCG package [36], with default values for all the parameters.

Acknowledgements

We thank R Fourme and J Perez for their assistance in data collection at LURE, Orsay, A van Dorsselaer for the mass spectroscopy results, J M Rondeau for useful discussions, T Simonson and A Varnek for help in theoretical electrostatic calculations, L Moulinier for assistance in the use of graphics, and R Ripp, S Uge and A Litt for computing assistance. This work was supported by the Centre National de la Recherche Scientifique (CNRS) through the UPR 9004, by the EU through the collaborative project CII-CT 93 0014, by the Institut National de la Santé et de la Recherche Médicale and the Centre Hospitalier Universitaire Régional, and by LIPHA industries through a contract with CNRS.

References

1. Larson, E.R., Lipinski, C.A. & Sarges, R. (1988). Medicinal chemistry of aldose reductase inhibitors. *Med. Res. Rev.* **8**, 159–186.
2. Wermuth, B. & Monder, C. (1983). Aldose and aldehyde reductase exhibit isocorticosteroid reductase activity. *Eur. J. Biochem.* **127**, 279–284.
3. Warren, J.C., Murdock, G.M., Ma, Y., Goodman, S.R. & Zimmer, W.E. (1993). Molecular cloning of testicular 20-α hydroxysteroid dehydrogenase. *Biochemistry* **32**, 1401–1406.
4. Branlant, G. (1982). Properties of an aldose reductase from pig lens. *Eur. J. Biochem.* **129**, 99–104.
5. Hoffman, P.L., Wermuth, B. & Von Wartburg, J.-P. (1980). Human brain aldehyde reductases: relationship to succinil semi aldehyde reductase and aldose reductase. *J. Neurochem.* **35**, 354–366.
6. Rondeau, J.-M., *et al.*, & Moras, D. (1992). Novel NADPH binding domain revealed by the crystal structure of aldose reductase. *Nature* **355**, 469–472.
7. Wilson, D.K., Bohren, K.M., Gabbay, K.H. & Quiocho F.A. (1992). An unlikely sugar substrate site in the 1.65 Å structure of the human aldose reductase holoenzyme implicated in diabetic complications. *Science* **257**, 81–84.
8. Bohrani, D.W., Harter, T.M. & Petrash, J.M. (1992). The crystal structure of the aldose reductase -NADPH binary complex. *J. Biol. Chem.* **267**, 24841–24847.
9. Tete, F., *et al.*, & Moras, D. (1995). Aldose Reductase from Pig Lens. *Eur. J. Med. Chem.* **30** S, 589s–603s.
10. Varma, S.D., Mizuno, A. & Kinoshita, J.H. (1977). Diabetic cataracts and flavonoids. *Science* **195**, 205–206.
11. Sarges, R., Bordner, J., Dominy, B.W., Peterson, M.J. & Whipple, E.B. (1985). Synthesis, absolute configuration, and conformation of the aldose reductase inhibitor sorbinil. *J. Med. Chem.* **28**, 1716–1720.
12. York, B.M. (1983). Spirotricyclic aromatic succinimide derivates as inhibitors of aldose reductase, European Patent Application 0092385.
13. Dvornik, D., *et al.*, & Merola, L. (1973). Polyol accumulation in galactodemic and diabetic rats: control by an aldose reductase inhibitor. *Science* **182**, 1146–1147.
14. Sestanj, K., *et al.*, & Dvornik, D. (1984). N-[5-(trifluoromethyl)-6-methoxy-1-naphthalenyl] thioxomethyl]- N- methylglycine (Tolrestat), a potent, orally active aldose reductase inhibitor. *J. Med. Chem.* **27**, 255–256.
15. Macleod, A.F., *et al.*, & Van Der Veen, E.A. (1992). A multicentre trial of the aldose-reductase inhibitor tolrestat, in patients with symptomatic diabetic peripheral neuropathy. *North European Tolrestat Study Group Diabetes Metab.* **18**, 14–20.
16. Stribling, D., Mirrlees, D.J., Harrison, H.E. & Earl, D.C.N. (1985). Properties of ICI 128,436, a novel aldose reductase inhibitor, and its effects on diabetic complications in the rat. *Metabolism* **34**, 336–344.
17. Sarges, R. & Oates, P.J. (1993). Aldose reductase inhibitors: recent developments. *Prog. Drug. Res.* **40**, 99–161.
18. Hotta, N., Kakuta, H., Ando, F. & Sakamoto, N. (1990). Current progress in clinical trials of aldose reductase inhibitors in Japan. *Exp. Eye Res.* **50**, 625–628.
19. Ward, W.H., *et al.*, & Howe, R. (1991). (2,6 dimethylphenylsulfonyl) nitromethane: a new structural type of aldose reductase inhibitor which follows biphasic kinetics and uses an allosteric binding site. *Biochem. Pharmacol.* **42**, 2115–2123.
20. Cook, P.N., *et al.*, & Howe, R. (1995). Kinetic characteristics of Zeneca ZD5522, a potent inhibitor of human and bovine lens aldose reductase. *Biochem. Pharmacol.* **49**, 1043–1049.
21. Wilson, D.K., Tarle, I., Petrash, J.M. & Quiocho, F.A. (1993). Refined 1.8 Å structure of human aldose reductase complexed with the potent inhibitor zopolrestat. *Proc. Natl. Acad. Sci. USA* **90**, 9847–9851.
22. Harrison, D.H., Bohern, K.M., Ringe, D., Petsko, G.A. & Gabbay, K.H. (1994). An anion binding site in human aldose reductase: mechanistic implications for the binding of citrate, cacodylate, and glucose 6-phosphate. *Biochemistry* **33**, 2011–2020.
23. Nakano, T. & Petrash, J.M. (1996). Kinetic and spectroscopic evidence for active site inhibition of human aldose reductase. *Biochemistry* **35**, 11196–11202.
24. Kubiseski, T.J., Hyndman, D.J., Morjana, N.A. & Flynn, T.G. (1992). Studies on pig muscle aldose reductase. Kinetic mechanism and evidence for slow conformational change upon coenzyme binding. *J. Biol. Chem.* **267**, 6510–6517.
25. Grimshaw, C.E., Shahbaz, M. & Putney, C.G. (1990). Mechanistic basis for non-linear kinetics of aldehyde reduction catalysed by aldose reductase. *Biochemistry* **29**, 9947–9955.
26. De Winter, H.L. & von Itzstein, M. (1995). Aldose reductase as a target for drug design: molecular modeling calculations on the binding of acyclic sugar substrates to the enzyme. *Biochemistry* **34**, 8299–8308.
27. Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489.
28. Ehrig, T., Bohren, K.M., Predergast, F.G. & Gabbay, K.H. (1994). Mechanism of aldose reductase inhibition: binding of NADP+/NADPH and alrestatin-like inhibitors. *Biochemistry* **33**, 7157–7165.
29. Barski, O.A., Gabbay, K.H., Grimshaw, C.E. & Bohren K.M. (1995). Mechanism of human aldehyde reductase: characterization of the active site pocket. *Biochemistry* **34**, 11264–11275.
30. Reymann, J.M., Rondeau, J.-M., Barth, P., Jacquinod, M., van Dorsselaer, A. & Biellmann, J.-F. (1992). Purification and electrospray mass spectrometry of aldose reductase from pig lens. *Biochem. Biophys. Acta* **1122**, 1–5.
31. Kabsch, W. (1988). Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J. Appl. Cryst.* **21**, 916–924.
32. Brünger, A.T. (1992). X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR. Yale University Press, New Haven, CT, USA.
33. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M.O. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
34. Urzhumtsev, A.G., Skovoroda, T.P. & Lunin, V.Yu. (1996). A procedure compatible with X-PLOR for the calculation of electron-density maps weighted using an R-free-likelihood based approach. *J. Appl. Cryst.* **29**, 741–744.
35. Kleywegt, G.J. & Jones, T.A. (1996). xdlMAPMAN and xdlDATAMAN – programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflections data sets. *Acta Cryst. D* **52**, 826–828.
36. Program Manual for the Wisconsin Package, Version 8, August 1994. 575 Science Drive, Madison, Wisconsin, USA 53711.
37. Evans, S.V. (1993). SETOR: hardware lighted three dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11**, 134–138.
38. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
39. Pavelites, J.J., Gao, J., Bash, P.A. & Mackerell, A.D. (1997). A molecular mechanics force field for NAD+, NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* **18**, 221–239.
40. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1991). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.