EUROPEAN JOURNAL OF

medicinal chemistry

Proceedings of the XIIIth International Symposium on Medicinal Chemistry

Editor: Jean-Claude Muller

Suppl to Vol 30 1995 – ISSN 0223-5234

ELSEVIER

Aldose Reductase from Pig Lens

F. Tête-Favier, P. Barth° , A. Mitchler, A. D. Podjarny, J.-M. Rondeau+, A. Urzhumtsev#, J.-F. Biellmann° & D. Moras

UPR de Biologie Structurale 9004, Institut de Génétique et de Biologie Moléculaire et Cellulaire du CNRS, 1 rue Laurent Fries, BP 163, 67404 Illkirch-Cedex, France.

° Laboratoire de Chimie Organique Biologique, Institut de Chimie, Université Louis Pasteur, 4 rue Blaise Pascal, 67008 Strasbourg-Cedex, France.

+ (Present address) Protein Crystallography Laboratory, Marion-Merrell-Dow Research Institute, Strasbourg Research Centre, 16 rue d'Ankara, 67080 Strasbourg Cedex.

#(Permanent address) IMPB RAN, Pushchino, Moscow Region, 142292, Russia.

Summary

Aldose Reductase (AR) is a NADPH-dependent enzyme implicated in long term diabetic complications. The protein exhibits the frequently observed $(\beta/\alpha)_8$ TIM barrel fold, a novel NADPH binding domain. Its crystal structure shows that the coenzyme is buried at the bottom of a deep hydrophobic cleft, and surrounded by a network of conserved hydrophilic residues. Extensive biochemical and mutagenesis studies confirm the crucial role of these residues in enzymatic activity.

A catalytic mechanism is presented, explaining the role of several conserved residues. H110 functions as the catalytic residue and Y48 stabilises an anion binding well which favours correct substrate orientation during catalysis. Two water channels link the active site, and in particular H110, with the bulk solvent and could provide the proton necessary for the reaction. Sequence comparisons show that these channels are likely to be present in all members of the aldo-ketoreductase superfamily.

1. Introduction

Although the physiological role of Aldose Reductase (AR) (EC 1.1.1.21) in healthy individuals is still a matter of debate, this 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 non-insulinodependent tissues [1]. In vitro, AR has been shown to act on a wide range of substrates, such as aldehydes, aldoses or corticosteroids, but it is most effective on steroid hormones [2-3]. This NADPH-dependent enzyme reduces a carbonyl group to a hydroxyl one, in an ordered bi-bi mechanism, where NADPH is bound first and $NADP^+$ released last. Upon formation of the enzyme-coenzyme-substrate ternary complex, a hydride is transferred from the C4 carbon atom of the nicotinamide ring (4-pro-R, *A*-face) of NADPH [4-5] to the carbonyl carbon of the substrate, while a proton is provided by the enzyme to the carbonyl oxygen. Like all other NAD(P)-dependent enzymatic reactions known to date, this reduction is stereospecific with respect to both substrate and coenzyme.

Aldose reductase belongs to a class of monomeric NADPH-dependent oxidoreductases of molecular weight in the range 35-40 kD, the aldo-ketoreductases [6] (Table 1). Published sequence data have shown that within this class aldose reductase and the frequently co-occurring aldehyde reductase have high sequence homology and therefore should fold in very similar three-dimensional structures [7-10]. In addition, data base searches have indicated strong sequence similarities between these enzymes and prostaglandin F synthase [11], 2,5-diketo-D-gluconate reductase [12], ρ-crystallin [13], chlordecone reductase [14] and a yeast protein encoded by the GCY gene [15].

Due to the therapeutic interest of AR, several crystallographic analyses have been performed on the enzyme from porcine, human and bovine sources [16-21]. These studies reveal highly similar structures of AR through species. The coenzyme is mostly buried, with the catalytic C4 atom barely accessible at the bottom of a deep pocket opened on one side of the enzyme. This pocket is clearly the substrate binding site, with the enzymatic activity taking place at its bottom. Based on the structure of the wild type enzyme, two residues were early proposed as the possible proton donors in the catalytic process : H110 and Y48 [17]. Although H110 is very favourably located, it was first discarded with preference for Y48, as the hydrophobic character of the cleft is expected to lower the pKa of H110 and therefore to favour the unprotonated form of the histidine, making it unsuitable as a proton donor.

Table 1: Sequence alignment of the members of the aldo-ketoreductase superfamily.
The sequences search and alignment were performed using the programs of the GCG package [28] using the SwissProt, GenEMBL and PIR-Prot data bases. The following sequences were included : soybean 6'-deoxychalcone synthase (6dcs-soy), barley AR (Aldr-hor), human chlordecone reductase (Chlr-hum), human dihydrodiol dehydrogenase (Hsdhd), bovine prostaglandin F synthase (Pgfs-bov), rat 3α-hydroxysteroid dehydrogenase (Didh-rat), frog ρ-crystalline (Cro-ranc), rat D(4)-3-ketosteroid reductase (Ksbr), bovine lens AR (Aldr-bov), pig lens AR [29] (Aldr-pig), human AR (Aldr-hum), kidney rabbit AR (Aldr-rab), rat lens AR (Aldr-rat), human aldehyde reductase (Aldx-hum), a protein coded by the GCY gene of baker's yeast (Gcy-yeas), a reductase from gene P100/11E of Leishmania major (P100-lei), D-sorbitol-6-phosphate dehydrogenase (S6pd-mal), xylose reductase from P. stipitis XYL1-gene (Psxyl1re) and the 2,5-diketogluconic acid reductase from Corynebacterium Sp. (2dkg-cor). The first line indicates the numbering used in the text, in multiples of 10, corresponding to the sequence of pig AR. The residues of groups I and II (see table 2) are indicated by + and x respectively. The last line indicates the strands (s) and the helices (h) as observed in the pig enzyme. Highlighted letters correspond to residues conserved in at least all the compared sequences but one.

2. X-Ray structures :

2.1 The enzyme structure :

The structure of apo AR from pig lens, initially solved at 2.5Å resolution in a triclinic crystal form with 4 monomers in the asymmetric unit [16,18], has now been refined at 2.0 Å resolution. The structure of the corresponding holoenzyme has been obtained at 2.8 Å resolution from a tetragonal crystal form with one monomer per asymmetric unit.

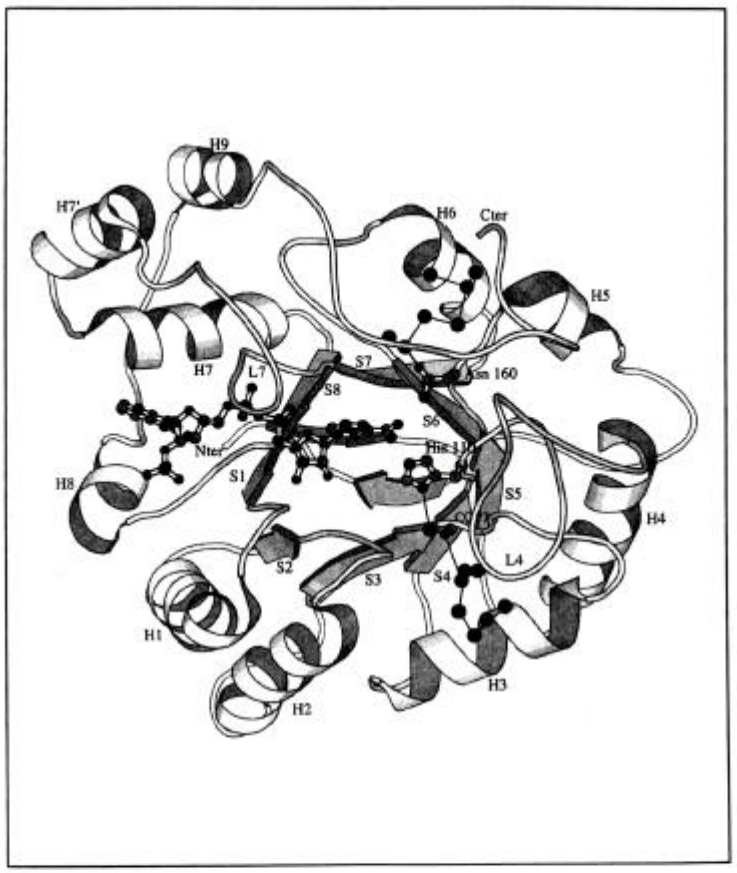

Figure 1: Overall schematic drawing of the holo form of Aldose Reductase, obtained with the program MOLSCRIPT [30].
Numbering of strands (S1, S2, ..., S8), helices (H1, H2, ..., H9), and loops (L4 and L7) is shown. $NADP^+$ (ball-and-stick) lies between strands 7 and 8, with loop L7 bent on it. The water molecules (balls) of the channels, connected by hydrogen bonds (dotted lines), link the outside of the protein to residues H110 and N160 (balls and sticks).

Pig lens aldose reductase comprises a single polypeptide chain of 314 aminoacids with an acetylated N-terminal serine. The enzyme is globular with a single structural domain, very similar to the archetypical structure of the triose phosphate isomerase $(\beta/\alpha)_8$ barrel [16] (Figure 1). It is legitimate to think that the above mentioned homologous proteins exhibit a similar fold.

2.2. Coenzyme binding :

$NADP^+$ lies in a very extended conformation at the C-terminal side of the strands (figure 1). Its binding involves numerous interactions similar to those reported for the human holoenzyme [17]. Significant conformational changes are observed : the most important one concerns loop 7 which spreads over the coenzyme. During this change, P222 and P225 switch from a trans peptide bond in the apoenzyme to a cis peptide bond in the holoenzyme. W20 and K21 occupy two different positions corresponding to the opening and the closing of a latch formed by the bulky aromatic residue and a salt bridge between K21 and D216. 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-catalysed reduction of D-glyceraldehyde [22,23].

2.3. The active site :

The nicotinamide ring of the coenzyme is largely buried at the bottom of a deep cleft with approximate dimensions of 4 x 15 Å wide and 15 Å deep, situated at the C-terminal side of the cylindrical sheet above strands S1 and S2. The only accessible atoms of $NADP^+$ are C5 and the hydride donor C4, defining thus the cleft as the only possible active site. The residues lining the cleft can be divided into 2 groups (table 2) :

• I) A mainly hydrophilic basis is constituted by residues D43, Y48, K77, H110, S159, N160, Q183 and Y209. Highly conserved in all the enzymes of the superfamily, they form a tight network of hydrogen bonds linked to the nicotinamide (figure 2a). S159, N160 and Q183, H-bonded to the amide group of the nicotinamide, assure the precise positioning of the nicotinamide ring for stereo selective transfer of the 4-pro-R hydrogen. Noteworthy also is the stacking interaction with Y209 which suggests stabilisation of the positively charged aromatic system formed upon oxidation of NADPH to $NADP^+$ (observed in the holoenzyme crystal form) by a charge-transfer complex (figure 2b).

amino acid	Contacts			Exposure to substrate	Sequence conservation
	H-Bond	V.d.Waals	Other		
D43	Q183 $NADP^+$		salt bridge K77	no	yes
Y48	K77	strand 1 strand 2		small	T in Cro
K77	Y48 CO 44	H110	salt bridge D43	no	yes
H110	H2O (Ch1)	W79 W111		medium hydrophilic	E in Ksbr
S159	$NADP^+$			no	C in ARhorvu
N160	$NADP^+$ Y209 H2O (Ch2)	W111		no	yes
Q183	D43 $NADP^+$			no	yes
Y209	N160		charge transfer complex $NADP^+$	small hydrophobic	aromatic conserved

Table 2a

amino acid	Contacts		Exposure to substrate	Sequence conservation
	H-Bond	V.d.Waals		
W20		W219 C298 $NADP^+$	large possible H-bond	aromatic conserved (T in Didh)
W79	residues other than groups 1,2	H110	medium hydrophobic	P in 2dkg
W111		W79 H110 N160	medium possible H-bond	aromatic conserved (S in Pgfs, M in Ksbr)
F122		W79 L300	large	no
P218		W219	medium	no
W219	residues other than groups 1,2	W20 P218 C298	large	no
C298		W219	small	no
L300		W111	small	no

Table 2b

Table 2 : The 2 groups of residues in the active site pocket of AR. a) Group I : Residues of the hydrophilic network around nicotinamide at the basis of the active site. Ch1 and Ch2 correspond to water chanels 1 and 2 respectively.b) Group II : Residues of the hydrophilic sides of the active site.

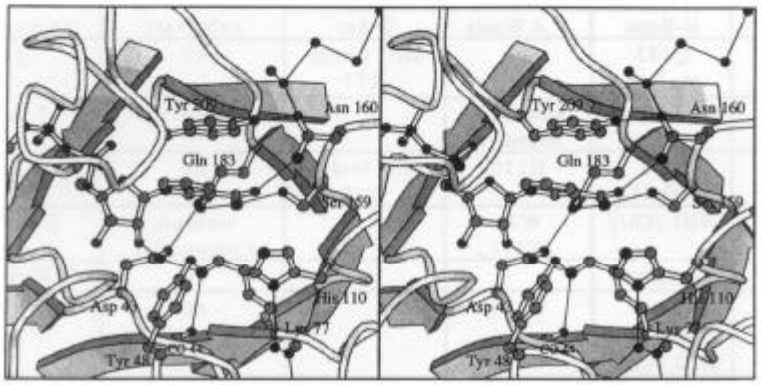

Figure 2a

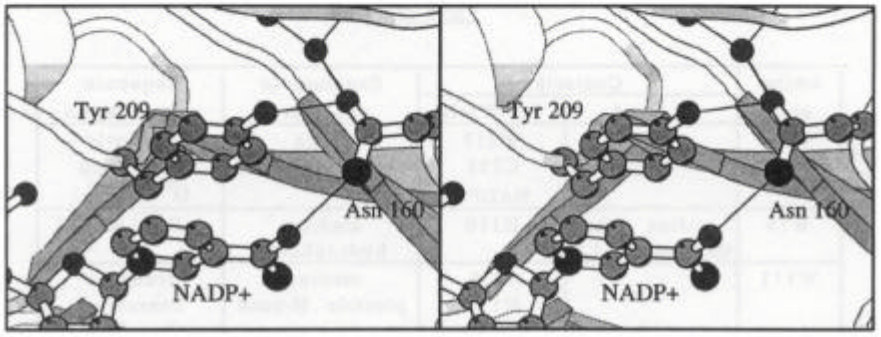

Figure 2b

Figure 2: Details of the active site of pig lens Aldose Reductase.

a) Residues of group I, $NADP^+$ and water molecules of the channels are shown in balls and sticks. The network of hydrogen bonds is shown in dotted lines. For clarity, helices and N-terminal parts of AR were omitted.

b) Stacking of Y209 and the nicotinamide ring. The view direction is perpendicular to the plane of the rings, which are 3.5 Å apart. Note that the arrangement of the rings corresponds to the geometry of a charge transfer complex.

• II) Above the coenzyme, residues W20, W79, W111, F122, P218, W219, C298, L300 form the walls of the hydrophobic cleft. W20, F122 and W219 face the active site, and are thus expected to make the major contacts with potential substrates (figure 3), as confirmed by the structure of the complex between Zopolrestat and the human holo AR [20]. In this group, the more the residues are near the outlets of the cleft, the less they are conserved.

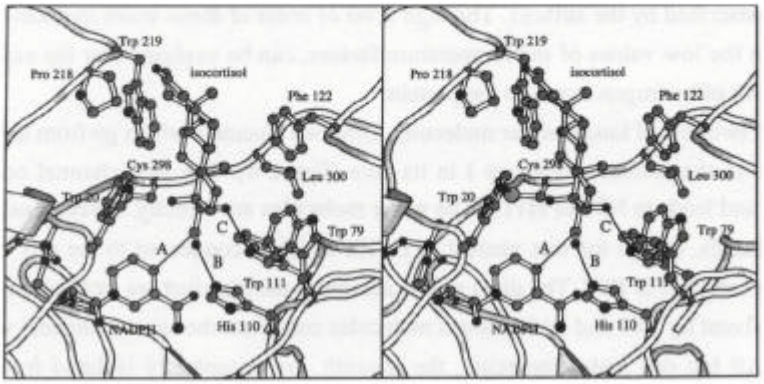

Figure 3 : Model of isocortisol fitted in the active site. AR is represented as a ribbon, isocortisol, $NADP^+$, H110, and residues of group II with balls and sticks. The thin lines show A) the catalytic contact (2.9 Å) between C21 of isocortisol and C4 of NADPH; B) the catalytic contact (3.3 Å) between O21 of socortisol and Nε2 of H110; C) the possible hydrogen bond (2.8 Å) between O20 of isocortisol and Nε1 of W111. The isocortisol model was derived from the crystal structure of corticosterone [31]. To dock it, the protein was kept rigid and a multidimensional search of optimal substrate-protein contacts was performed varying the position and orientation of the substrate and its dihedral angles C17-C20 and C20-C21 in the range allowed by substrate-substrate contacts. During this search, the substrate positions with distances between O21 and Nε1 of H110 and between C21 and C4 of the coenzyme larger than 4Å or with bad steric contacts were excluded. For the solutions of this search, the dihedral angles C17-C20 and C20-C21 were varied to check also possible contacts with Y48.

2.3. Two water channels :

Among the water molecules localised in the triclinic structure of the pig lens apo enzyme, two sets are highly conserved in the four monomers which constitute the asymmetric unit of the crystal. A difference electron density map (F_{obs} - F_{cal}, α_{cal}) clearly shows the density corresponding to the same water molecules in the tetragonal holoenzyme crystals. A strong independent evidence for the conservation of these water molecules is provided by the crystal structure at 1.65 Å of the human holoenzyme [17]. Despite the differences in source, crystallisation conditions and space group, this structure exhibits these water molecules at the same locations. These molecules have been reported to the PDB (entry 1ADS) without being explicitly described by the authors. The high level of order of these water molecules, which is reflected in the low values of the temperature factors, can be explained by the existence of a tight network of hydrogen bonds to the protein.

The two sets of linked water molecules form two channels which go from the surface of the protein to the residues of group I in its core (figure 4). The first channel consists of 8 molecules and leads to Nδ1 of H110. The water molecules are directly linked to each other by hydrogen bonds, except the one nearest to H110, which is connected to the rest through the main chain carbonyl of K77. The three waters nearest to the histidine are totally buried, isolated from the solvent by V47 and P121. Seven molecules constitute the second channel which leads to N160. All but one water molecule, the seventh, are completely isolated from the bulk solvent, mainly by residues A299 and R296.

3. Discussion

3.1. Functional implications

The presence of the two water channels could be particularly relevant to the proton transfer step involved in the catalytic mechanism of this class of enzymes. The proton transferred from the enzyme to the substrate is taken from one of the residues of group I, which embed the nicotinamide ring in a hydrophilic environment at the bottom of a hydrophobic active site. Among the residues neighbouring the coenzyme, a natural candidate for providing the proton is H110, which had originally been discarded due to its hydrophobic environment. The discovery of the hydrophilic link leading to H110 sheds a new light on the structural environment of this crucial residue and undermines the former argument. Indeed, this water channel allows stabilisation of the protonated form of H110 and facilitates the exchange of protons with the solvent.

Fig.4a

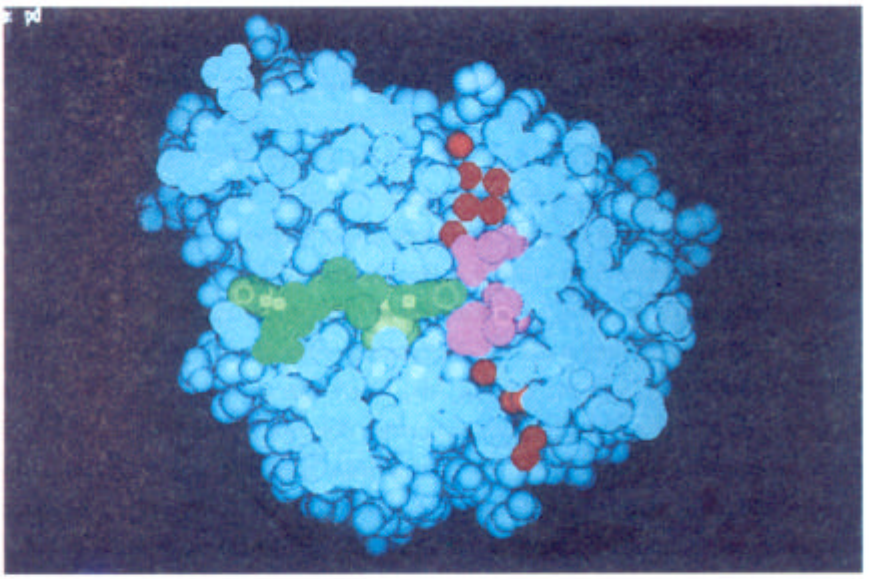

Fig.4b

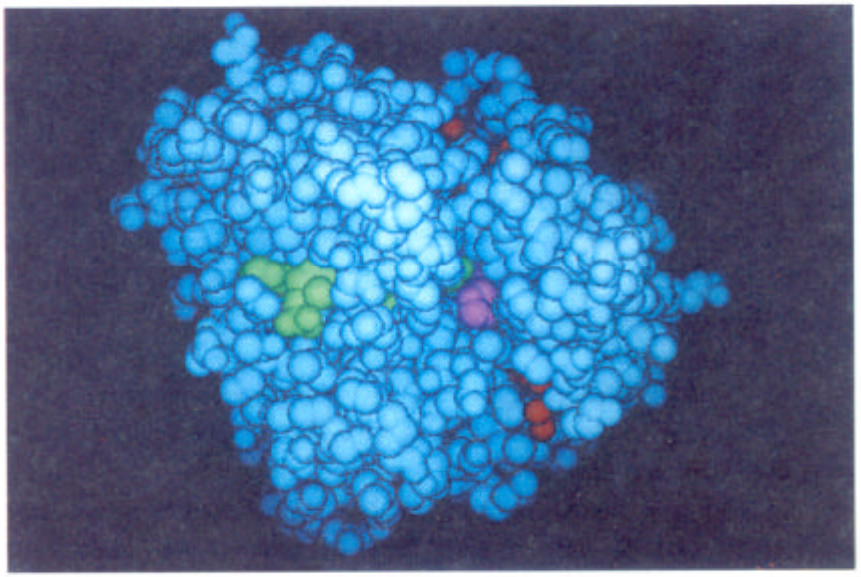

Figure 4: The water channels in Aldose Reductase.**a)** Space-filling model, viewed along the barrel axis. $NADP^+$ (green) is buried in the protein, the only accessible atoms of the nicotinamide ring being C5 and the hydrogen donor C4. Note the accessibility of H110 (magenta) and the outlet of the buried water channels (red). Since the channel leading to N160 (top channel) is almost totally buried, an additional water molecule, fully exposed, has been highlighted in red in order to mark the entry. Other parts of the complex are in light blue. **b)** Cut view of the same space-filling model than in 4a. The full water channels in red lead to the magenta residues H110 (bottom) and N160 (top). $NADP^+$ is in green, and the other parts of the complex are in light blue.

The role of the second water channel leading to N160 is less clear. It should be noted that, although N160 directly interacts with the nicotinamide, and is totally conserved through the aldo-ketoreductase superfamily, no mutagenesis experiment was never made for this residue. This channel role could be to shield the charge of the salt bridge betwen R296 and E185 from the rest of the molecule. It could also allow water pressure equilibration during enzyme turnover, as both binding of substrates and release of products implies a concomitant flow of water molecules from or to the active site cleft [24].

The analysis of activity as a function of pH does not resolve the question of the proton donor : even if the optimum pH for the substrate reduction strongly accounts for H110 (see [25] for bovine AR, and [4] for pig AR, where the optimum pH is 4.8 for a aldehydic substrate, which is an unlikely value for proton donation by the hydroxyl group of Y48), the local environment of a residue can modify its pKa, leading to uncertain conclusions. Clearly, site-directed mutagenesis experiments aimed at blocking either or both water channels would be extremely welcome to ascertain their role in the mechanism of aldo-ketoreductases.

3.2.The anionic site :

In addition to the water channels, another remarkable feature of the active site of AR is the existence of an anion binding well constituted with D43, Y48, K77, and the nicotinamide ring of $NADP^+$ [26]. Experimental evidence for the existence of this anion well was provided by the identification of a previously unexplained electron density in the active site of the holoenzyme as a citrate ion. The bound citrate can be replaced in the crystal by cacodylate or glucose 6-phosphate. Analysis of electrostatic potentials in pig lens AR confirm the presence of this anionic site. The enzyme can thus accommodate the negatively charged transition state that appears during substrate reduction.

3.3. Proposed mechanism and mutagenesis experiments :

Altogether, the structural features of AR active site support the following views on the enzymatic reduction of aldehydes by AR:

1 Upon formation of the ternary complex, polarisation of the carbonyl bond of the substrate through hydrogen bonding to H110 Nε2 facilitates the nucleophilic attack of the carbonyl atom by the hydride originating from the nicotinamide C4 atom.

2 The positively charged aromatic system on the nicotinamide ring is stabilised by charge-transfer interactions with Y209.

3 The anion well favours the formation of the negatively charged alkoxide.

4 A proton is transferred from the Nε2 of H110 to the alkoxide oxygen.

Site directed mutagenesis experiments concern residues D43, Y48, K77 and H110 in group I aminoacids [21,27]. Their modification considerably affects the enzyme activity. The very low residual activities are differently perceived by the authors : for example, H110N, with a catalytic efficiency Kcat/Km of 0.00009% of that of the wild type, is still active in [27], while Y48S is not, with 0.0003% in [21]. Furthermore, the three dimensional structure of the Y48H mutant of human AR revealed that additive water molecules can replace the functional group of the mutated residue, preventing unambiguous conclusions. Thus from site-directed mutagenesis combined with kinetic and crystallographic studies, Bohren *et al.* [21] proposed that Y48 is the proton donor and H110 directs stereochemical selectivity in the reduction of the human AR.

Howewer H110 cannot be excluded as the proton donor, since the water channels invalidate the argument of an unprotonated state of the residue at the bottom of a hydrophobic cleft. Furthermore, this hypothesis fits the site-directed mutagenesis data which show that the modification of either the proton donor (H110) or any of the amino-acids involved in the formation of the anion binding site (as Y48, D43, K77...) considerably affects the enzyme activity, and even suppresses it (depending on the choice of the replacing residue, and the potential surrogate role of water molecules).

4. Conclusion

The complete enzymatic mechanism of AR is still uncertain but several points are now well established : NADPH binds first to the enzyme and induces some structural changes, in particular the movement of a loop which completes the arrangement of the substrate binding site. Subsequently the ternary complex is formed, the hydride of the nicotinamide ring is transferred to substrate, with $NADP^+$ stabilised by a charge transfer complex with Y209. Then a proton is delivered to the substrate, by either Y48 or H110. This second amino-acid is the best candidate, since its protonated state is assured by one of the water channels observed in the structure of AR.

To verify the agreement of the proposed active site structure with the known substrate specificity of Aldose reductase, 20-α-isocortisol was docked in the active site (figure 3). The study showed a position that correctly places the atoms for reaction, consistent with H110 as the proton donor. It emphasised the capability of the bottom of the cleft (W20 and W111) to maintain a substrate in the active site, while the larger top of the cleft accommodates the shape and the properties of the different substrates.

Although some drugs designed to inhibit AR already exist, the conception of more efficient molecules still constitutes a major pharmaceutical aim. AR possess a large substrate specificity; furthermore, it belongs to the aldo-ketoreductase superfamily, which contains

several highly homologous enzymes. This explains the difficulty to conceive specific inhibitors. The specificity of AR comes from residues located on loops of variable length probably liable to move, leading to difficulties in making use of them in the rational conception of inhibitors.

References

1. Dvornik D. (1987). In *Aldose Reductase Inhibition*. McGraw-Hill, New York.
2. Wermuth B. (1985). In *Enzymology of Carbonyl Metabolism 2*. pp. 209-230, Alan R. Liss Inc.
3. Warren J. C., Murdock G. M., Ma Y., Goodman S. R. & Zimmer W. E. (1993). *Biochemistry* **32**, 1401-1406.
4. Branlant G. (1982). *Eur. J. Biochem.* **129**, 99-104.
5. Hoffman P.L., Wermuth B, & Von Wartburg J.-P. (1980). *J. Neurochem.* **35**, 354-366.
6. Wermuth B. (1985). *Prog. clin. biol. Res.* **174**, 209-230.
7. Bohren K.M., Bullock B., Wermuth B. & Gabbay K.H. (1989). *J.Biol.Chem.* **264**, 9547-9551.
8. Chung S. & LaMendola J. (1989). *J.Biol.Chem.* **264**, 14775-14777.
9. Nishimura C., Wistow G. & Carper D. (1989). *Prog. clin. biol. Res.* **290**, 211-220.
10. Schade S.Z., Early S.L., Williams T.R., Kézdy F.J., Heinrikson R.L., Grimshaw C.E. & Doughty C.C. (1990). *J.biol.Chem.* **265**, 3628-3635.
11. Watanabe K., Fujii Y., Nakayama K., Ohkubo H., Kuramitsu S., Kagamiyama H., Nakanishi S. & Hayaishi O. (1988). *Proc.natl.Acad.Sci.USA* **85**, 11-15.
12. Anderson S., Marks C. B., Lazarus R., Miller J., Stafford K., Seymour J., Light D., Rastetter W. & Estell D. (1985). *Science* **230**, 144-149.
13. Fujii Y., Watanabe K., Hayashi H., Urade Y., Kuramitsu S., Kagamiyama H. & Hayaishi O. (1990). *J.biol.Chem.* **265**, 9914-9923.
14. Winters, C.J., Molowa, D.T. & Guzelian, P.S. (1990). *Biochemistry* **29**, 1080-1087.
15. Oechsner, U., Magdolen, V. & Bandlow, W. (1988). *FEBS Lett.* **238**, 123-128.
16. Rondeau J.-M., Tête-Favier F., Podjarny A., Reymann J.-M., Barth P., Biellmann J.-F. & Moras D. (1992). *Nature* **355**, 469-472.
17. Wilson D. K., Bohren K. M., Gabbay K. H., & Quiocho F. A (1992). *Science* **257**, 81-84.
18. Tête-Favier F., Rondeau J.-M., Podjarny A. & Moras D. (1993). *Acta Cryst.* **D49**, 246-256.

19. Bohrani D. W., Harter T. M., & Petrash J. M. (1992). *J. Biol. Chem.* **267**, 24841-24847.
20. Wilson D. K., Tarle I., Petrash J. M., & Quiocho F. A. (1993). *Proc. Natl. Acad. Sci.* **90**, 9847-9851.
21. Bohren K. M., Grimshaw C. E., Lai C.-J., Harrison D.H., Ringe D., Petsko G.A. & Gabbay K. H. (1994). *Biochemistry* **33**, 2021-2032.
22. Kubiseski T.J., Hyndman D.J., Morjana N.A., & Flynn T.G. (1992). *J. Biol. Chem.* **267**, 6510-6517.
23. Grimshaw C.E., Shahbaz M., & Putney C.G. (1990). *Biochem.* **29**, 9947-9955.
24. Meyer E. (1992). *Protein Sciences* **1**, 1543-1562.
25. Liu S.-Q., Bhatnagar A., & Srivastava S. K. (1993). *J. Biol. Chem.* **268**, 25494-25499.
26. Harrison D. H., Bohren K. M., Ringe D., Petsko G. A. & Gabbay K. H. (1994). *Biochemistry* **33**, 2011-2020.
27. Tarle I., Borhani D. W., Wilson D. K., Quiocho A., & Petrash J. M. (1993). *J. Biol. Chem.* **268**, 25687-25693.
28 Genetics Computer Group (1991). Program Manual for the GCG Package, Version 7, April 1991. 5875 Science Drive, Madison, Wisconsin, USA 53711.
29 Jaquinod M., Potier N., Klarskov K., Reymann J.-M., Sorokine O., Kieffer S., Barth P., Andriantomanga V., Biellmann J.-F. & Van Dorsselaer A. (1993). *Eur.J.Biochem* **218**, 893-903.
30. Kraulis P. J. (1991). *J. Appl. Cryst.* **24**, 946-950.
31 Campstein H., Dupont L., Dideberg O., & Mandel N. (1973). *Acta Cryst.* **B29**, 1726 -1728.