Structure-activity relationships of human AKR-type oxidoreductases involved in bile acid synthesis: AKR1D1, AKR1C4

Wen Hwa Lee, Petra Lukacik, Kunde Guo, Emilie Ugochukwu, Kathryn L. Kavanagh, Brian Marsden, Udo Oppermann

► To cite this version:

Wen Hwa Lee, Petra Lukacik, Kunde Guo, Emilie Ugochukwu, Kathryn L. Kavanagh, et al.. Structure-activity relationships of human AKR-type oxidoreductases involved in bile acid synthesis: AKR1D1, AKR1C4. *Molecular and Cellular Endocrinology*, 2009, 301 (1-2), pp.199. 10.1016/j.mce.2008.09.042 . hal-00532092

HAL Id: hal-00532092

<https://hal.science/hal-00532092>

Submitted on 4 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Title: Structure-activity relationships of human AKR-type oxidoreductases involved in bile acid synthesis: AKR1D1, AKR1C4

Authors: Wen Hwa Lee, Petra Lukacik, Kunde Guo, Emilie Ugochukwu, Kathryn L. Kavanagh, Brian Marsden, Udo Oppermann

PII: S0303-7207(08)00447-4
DOI: doi:10.1016/j.mce.2008.09.042
Reference: MCE 7006

To appear in: *Molecular and Cellular Endocrinology*

Received date: 1-7-2008
Revised date: 27-9-2008
Accepted date: 27-9-2008

Please cite this article as: Lee, W.H., Lukacik, P., Guo, K., Ugochukwu, E., Kavanagh, K.L., Marsden, B., Oppermann, U., Structure-activity relationships of human AKR-type oxidoreductases involved in bile acid synthesis: AKR1D1, AKR1C4., *Molecular and Cellular Endocrinology* (2008), doi:10.1016/j.mce.2008.09.042

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Structure-activity relationships of human AKR-type oxidoreductases
involved in bile acid synthesis: AKR1D1 and AKR1C4.**

Wen Hwa Lee¹, Petra Lukacik¹, Kunde Guo¹, Emilie Ugochukwu¹, Kathryn L
Kavanagh¹, Brian Marsden¹, Udo Oppermann^{1,2}

¹Structural Genomics Consortium, University of Oxford, Oxford OX3 7LD, UK

²Botnar Research Centre, Oxford Biomedical Research Unit, OX3 7LD, UK.

Keywords: aldo-keto reductase, bile acid metabolism, homology modelling, docking,
X-ray crystallography

Abbreviations: AKR: aldo-keto reductase; SDR: short-chain
dehydrogenase/reductase:

Author for correspondence:

Wen Hwa Lee

SGC, ORCRB - Old Road Campus, Roosevelt Drive, Oxford, UK - OX3 7DQ

e-mail: wenhwa.lee@sgc.ox.ac.uk

Phone: +44 1865 617577

Fax: +44 1865 617575

ABSTRACT

Two members of the human aldo-keto reductase (AKR) superfamily participate in the biosynthesis of bile acids by catalyzing the NADP(H) dependent reduction of 3-keto groups (AKR1C4) and Δ^4 double bonds (AKR1D1) of oxysterol precursors.

Structure determination of human AKR1C4 and homology modelling of AKR1D1 followed by docking experiments were used to explore active site geometries. Substrate docking resulted in ligand poses satisfying catalytic constraints, and indicates a critical role for Trp227/230 in positioning the substrate in a catalytically competent orientation.

Based on the evidence gathered from our docking experiments and experimental structures, this tryptophan residue emerges as a major determinant governing substrate specificity of a subset of enzymes belonging to the AKR1 subfamily.

INTRODUCTION

Bile acid synthesis and excretion comprise the major pathway of cholesterol catabolism and elimination in mammals. Furthermore, this conversion adds the emulsifying property to the bile acids, which is utilized in the body to solubilise hydrophobic molecules obtained from the diet (e.g. cholesterol, lipids and lipophilic vitamins) to allow absorption in the digestive tract.

The pathway from cholesterol to bile acids requires at least 17 distinct enzymatic steps, with a critical involvement of several members of the Cytochrome P450 family, as well as oxidoreductases of the aldo-keto reductase (AKR) and short-chain dehydrogenase/ reductase (SDR) families (Penning 1999; Oppermann, Filling et al. 2003; Russell 2003; Bauman, Steckelbroeck et al. 2004; Jin and Penning 2007). Among these, two human AKR enzymes, Δ^4 -3oxosteroid, 5 β - reductase (AKR1D1) and 3 α -hydroxysteroid dehydrogenase (AKR1C4) carry out important modification reactions in the A-ring of the bile acid intermediates (Russell 2003). The first enzyme, AKR1D1 catalyzes the NADPH-dependent reduction of the double bond in the steroid A-ring, which introduces a 90 degrees bend between rings A and B. The second catalyst, AKR1C4 completes the ring modification through the NADPH-dependent reduction of the 3-oxo group to an alcohol in the stereochemical alpha configuration.

A central intermediate in the bile acid pathway is 4-cholesten-7 α -ol-3one, (Russell 2003), which can be directly processed by AKR1D1 and AKR1C4 respectively, resulting in the synthesis of chenodeoxycholic acid. Alternatively it can be modified by sterol 12 α -hydroxylase (CYP8B1) prior to the AKR1D1 and AKR1C4 steps, leading to cholic acid as the final product. The fact that both AKR1D1 and AKR1C4 can process these two types of intermediates illustrates the well-known ability of the AKRs to process a wide range of compounds, including xenobiotic substrates (Jez and Penning 2001). Moreover, at least for AKR1D1 a physiological role through inactivation of 3-oxo, Δ^4 steroid hormones (progestins, glucocorticoids, androgens, mineralocorticoids) has been demonstrated (Palermo, Marazzi et al. 2008).

This versatility and plasticity in the AKR family is achieved by a rather flexible active site, which is composed by three loops (namely A, B and C). These loops are not highly conserved among the different members of the AKR family, as expected for these promiscuous enzymes. Interestingly, most of the family 1 AKR structures solved so far have shown that loops A, B and C adopt similar conformations –

independent of the presence or the nature of ligands in these structures. This suggests that despite the variability in sequence, the folds observed for these loops (AKR family 1) are required for proper binding of the substrate. This would indicate that substrate promiscuity can be explained just by minor side-chain rearrangements instead of major loop conformational changes.

While the active-site loops are variable, the catalytic machinery is well conserved in all AKRs and comprises Lys48, Asp50, Tyr55 and His117 (numbering from rat AKR1C9 sequence). A notable exception is AKR1D1, where His117 is replaced by a Glu residue to provide an alternative mode of catalysis (Jez and Penning 1998).

In order to understand structure-function relationships of AKR-type bile acid oxidoreductases we used X-ray crystallography, homology modelling and *in silico* docking to compare their molecular properties.

MATERIAL AND METHODS

Protein expression and purification: An N-terminal His6-tagged expression construct cloned into the IPTG inducible pNIC28-Bsa4 vector was expressed in TB medium and purified to apparent homogeneity by immobilized metal affinity (IMAC) and size-exclusion chromatography. Prior to crystallization, the His tag was removed by His-tagged TEV protease treatment and subsequent purification by IMAC.

Crystallisation of AKR1C4 and data collection: Crystals were grown by vapor diffusion at 4°C in 150 nl sitting drops. NADPH at a final concentration of 5 mM was added to the protein just prior to crystallisation. The drops were prepared by mixing 75 nl of protein solution and 75 nl of precipitant consisting of 2M tri-potassium citrate. Crystals were transferred to a cryo-protectant consisting of 20% glycerol, 80% well solution before flash-cooling in liquid nitrogen. Crystals diffracted to 2.4 Å and a dataset was collected with an in-house rotating anode generator (Rigaku FR-E superbright). Structure determination was achieved through molecular replacement. Data collection and refinement statistics are given in Table 1. The atomic coordinates and structure factors were deposited in the PDB (accession code 2FVL).

Homology modelling

In silico docking and homology modelling was performed using the ICM software package (Abagyan, Totrov et al. 1994). In order to perform docking experiments, homology modelling procedures were necessary in two instances. The first one was performed for the binary cofactor complex of AKR1C4 to render the structure suitable

for docking, since the entrance to the active site was occluded by Trp227 and loop B in the crystal structure. We have identified the flexing portion of loop B (Gly220 to Val234) and then built homology models using the sequence from AKR1C4 and two templates (see below). From the resulting homology models, we extracted only the flexing portion of loop B and grafted it back to the deposited crystal structure of AKR1C4 (2FVL) – thus replacing the closed loop B by an opened form.

The second modelling procedure was performed for the full-length AKR1D1, since no crystal structures were available at the inception of this project. For both modelling procedures, two templates were used – human AKR1C2 (PDB code 1IHI) for its high identity level (58% to human AKR1D1 and 82% to human AKR1C4) to the sequences to be modelled and rat AKR1C9 (PDB code 1AFS) for presenting a steroid compound bound in a productive pose (identity levels: 54% to human AKR1D1 and 69% to human AKR1C4). Homology modelling and global minimisation to solve potential clashes derived from the method (e.g. substitutions, loop modelling/grafting) as well as hydrogen bonding network optimisation were performed using the protocols implemented in the program ICM (Cardozo, Totrov et al. 1995).

The resulting models were used as receptors for *in silico* docking with their respective substrates and products (4-cholesten-7 α -ol-3-one and 5 β -cholestan-7 α -ol-3-one for AKR1D1 and 5 β -cholestan-7 α -ol-3-one and 5 β -cholestan-3 α ,7 α -diol for AKR1C4). Docking was performed with fully flexible ligands and potential grid maps representing five different features of the receptor as implemented in the program ICM (Totrov and Abagyan 1997). No restraints were introduced to either receptor or ligand to allow full exploration of the active site. The best poses were visually screened and selected if distance requirements for catalysis were found to be acceptable for the atoms involved in the proposed mechanisms – i.e. with the 3-keto/hydroxyl group positioned within interaction distance of the catalytic Tyr55 (2.5 ~ 3.5 Å) and with the carbons to receive/ donate hydrides positioned within acceptable distance (3.5 ~ 4.5 Å) of the donor/ receptor C4 atom of the nicotinamide ring of the NADP(H)cofactor.

RESULTS

We have prepared two sets of models that are suitable for docking experiments. Each set is based on a structural template and is comprised of (i.) an ‘opened’ version of the

crystal structure of AKR1C4 (2 FVL) with a remodelled loop B and (ii.) a complete homology model of AKR1D1. The first set was built based on the structure of human AKR1C2 (1IHI), hereafter called “set-1IHI” and the second set was built using the structure of rat AKR1C9 (1AFS) as template, called “set-1AFS”.

The two sets of models were then used as receptors in unrestrained docking experiments where substrate and product of each enzyme were used as ligands. The results showed that poses satisfying the geometric constraints necessary for catalysis were only found when models from the set-1AFS were used as receptors.

As such, the best pose found for the substrate of AKR1D1 4-cholesten-7 α -ol-3-one positions the carbonyl group close to the catalytic residues Glu120 and Tyr58 and at the same time places the C5 in an ideal position for hydride transfer (Fig. 1A). On the receptor side, the indole ring of Trp230 is found swung inwards, towards the cofactor, whilst Tyr26 coordinates the 7 α -hydroxyl group.

Docking of the product of AKR1D1 suggests that the active site can accommodate the bend in the steroid ring system introduced by 5 β reduction (Fig. 1B). Tyr26 continues to engage the 7 α -hydroxyl, but the reduced 5 β together with the bend repositions the carbonyl group away from the Glu120. Concomitantly, Tyr132 moves out towards the solvent allowing the D-ring and the carbon side-chain to be raised as a consequence of the bending (not shown).

The results of docking both substrate and product into the set-1AFS model of AKR1C4 suggest that no significant change is to be expected from either the ligand or the receptor side of the system (Fig. 1C, D). Either one of the ligands could be docked in a pose satisfying the catalysis constraints, with the C3 carbonyl/hydroxyl group positioned within non-covalent bonding distance of the catalytic residues His117 and Tyr55. C3 was also positioned in a distance and geometry favourable for hydride transfer to/from the C4 of the nicotinamide ring.

Docking results obtained using models from the set-1IHI as receptors have generated poses that are positioned mostly across the active site – i.e. in an intermediate pose between parallel and orthogonal to the cofactor molecule, but still not satisfying geometric requirements for catalysis. Also, these poses are mostly found to be upside-down, with the C18 and C19 methyl groups pointing away from the Trp230 indole ring. (Fig. 1E).

DISCUSSION

Docking results using different models suggest that the ligand binding mode in AKR1C4 and AKR1D1 is strongly correlated with the positioning of Trp227/230. Ligand docking that resulted in poses that are consistent with catalysis mechanism on positions 3 and 5 (ring A) could only be achieved for models incorporating a side-chain rotamer of Trp227/230 pointing towards the core of the enzyme.

Interestingly, when we revisited several previously solved AKR structures (AKR1C1 (1MRQ), AKR1C2 (1IHI, 1J96), AKR1C3 (1RY0, 1S1P), AKR1C4 (2FVL), rat AKR1C9 (1AFS) and lately AKR1D1 (3BUV, 3BUR, 3COT, 3BV7, 3CMF)) it was clear that the conformation of the side chain of Trp227/230 could be clustered into two rotamers: The first has the indole positioned across the entrance of the active site (e.g. as seen in 1IHI) whereas the second one had the indole pointing inwards to the core of the protein (e.g. as seen in 1AFS) – thus resembling the two states of a switch. We will hereafter refer to Trp227/230 and the equivalent tryptophan residues in other members of the same family as “switch tryptophan”.

In several AKR structures in complex with steroids we can find the switch tryptophan adopting the ‘across’ rotamer (e.g. AKR1C1 (1MRQ), AKR1C2 (1IHI), AKR1C3 (1RY0); Fig. 2A). The indole ring seems to serve as a platform, favouring stacking interactions with steroid molecules. When ring stacking occurs with the less bulky alpha side of the steroids, the C18 and C19 methyl groups are forcefully oriented away from the indole rings. This is indeed observed in the crystal structures of several AKR enzymes, including human AKR1C2 (1IHI: UDCA; 1J96: testosterone) and AKR1D1 (3BUR: testosterone). From these observations we conclude that the ‘across’ orientation is the favoured rotamer, even in the absence of a steroid ligand.

In the rat AKR1C9 (1AFS: testosterone) and the recent human AKR1D1 (3COT: progesterone; 3CMF: cortisone), the structures in complex with a steroid have the switch tryptophan adopting the ‘inward’ rotamer. In this scenario the beta side of the steroid is facing down, with the bulky C18- and C19-methyl groups occupying an aliphatic cleft where the side walls are formed by the switch tryptophan itself and another hydrophobic residue (Fig. 2B).

Importantly, the rotamer adopted by the switch tryptophan can define activity towards positions 3, 17 and 20 of a steroid. A structural exception to the trend described above

is found in the structure of human AKR1C1 in complex with progesterone (1MRQ). In this structure the switch tryptophan adopts an ‘across’ rotamer and yet the steroid is bound with the beta side facing towards the platform presented by the indole rings (Fig. 2C). This is possible, however, when the steroid is bound in an inverted pose, with the ring D positioned closer to the cofactor, instead of the ring A. In this conformation, the 20-keto group can be seen near the C4 of the nicotinamide ring (4.1 Å to C20) – therefore consistent with the 20a-HSD activity displayed by AKR1C1.

During the course of this study, five structures of human AKR1D1 (PDB codes 3BUV, 3BUR, 3COT, 3BV7 and 3CMF) in were deposited in the PDB (Di Costanzo, Drury et al. 2008) - three of these structures were solved in complex with steroid compounds, providing an excellent opportunity to validate our models and docking results.

An overall structure comparison of the experimental structures of AKR1D1 and our homology model derived from rat AKR1C9 (1AFS) shows that both are rather similar, including the general trace adopted by the loops A, B and C, despite the fact that these are positioned differently to some extent (Fig. 3C). Importantly, the most remarkable difference in the active site between the five experimental structures of AKR1D1 is the side chain of the switch tryptophan. Two types of rotamers are found, according to the presence and type of steroids in the active site.

In both AKR1D1 structures where the steroid is bound in a competent pose (3COT: progesterone and 3CMF: cortisone) i.e. with satisfying distances and geometry for 5 β -reductase activity, the switch tryptophan adopts the ‘inwards’ rotamer, as we predicted with the docking experiment using our homology model of AKR1D1 and its substrate.

In the remaining three AKR1D1 structures (3BUV, 3BUR, 3BV7) the switch tryptophan adopts the ‘across’ rotamer. Two of these structures do not have a steroid or any ligand bound in the active site (3BUV and 3BV7), while one has testosterone bound in it (3BUR). However in AKR1D1 (3BUR), testosterone is bound with the beta side facing away from the indole at the same time that its longer axis along the four rings is rotated 90 degrees in the same plane of the rings – i.e. also ‘across’ the pocket. Again, these results validate our docking experiments using the homology model of an ‘across’ rotamer (i.e. set-1IHI) where the poses were all non-productive

and cluster roughly along the same pose of the testosterone found in 3BUR. This can be explained by the spacious pocket formed by loop B that exists when the switch tryptophan adopts the ‘across’ rotamer (Fig. 1F). In the structures of AKR1C2 (PDB code 1IHI) and AKR1C1 (PDB code 1MRQ) a bulky Histidine residue involved in coordination of the cofactor occupies this pocket, whereas in AKR1D1 and AKR1C4 the same position is occupied by a smaller serine. A summary of the poses adopted by steroids bound to AKR enzymes (crystallographic structures and docking experiments) and the side-chain rotamer adopted by the switch tryptophan in each case can be found on table 2.

Interestingly, a previous docking study has pointed out that the only difference observed between the substrate binding pockets of AKR1C1 (1MRQ) and AKR1C2 (1IHI) – namely the substitution of Leu54 by a Valine – was sufficient to change the 20 α -HSD activity of AKR1C1 to the 3 β -HSD activity of AKR1C2 (Jin and Penning 2006). The same residue (Leu/Val54) had previously been identified in another study employing molecular docking as the main reason for the differential reactions of the same substrate (5 α -DHT) by AKR1C1 (3 β -HSD reaction) and AKR1C2 (3 α -HSD reaction)(Steckelbroeck et al. 2004). An observation from the docking experiments by Jin and Penning is that the narrowest point of the substrate pocket is limited by the Leu54 and the Trp227 (the equivalent of the aforementioned ‘switch tryptophan’) therefore forcing the ligand to bind with its beta face towards the Trp227. The substitution of Leu54 by a less bulky Valine residue would reshape the pocket in order to allow the ligand to flip in the ring plane and present the alpha face to the Trp227 instead. As mentioned before, the Trp227 residue in AKR1C1 and AKR1C2 cannot swing inwards the pocket where the bulky His222 is positioned. Therefore the Leu54Val substitution is probably an alternative to the swinging Trp227 to reshape the pocket and remove stereochemical constraints necessary for a specific catalysis, assuming that the loops do not undergo major rearrangements.

These findings suggest that the plasticity of the substrate pocket can be fine-tuned with small rearrangements of the side-chains, with the degree of freedom dependent on the residue composition of the pocket. Hence, while for some cases there will be sufficient volume for the manoeuvring of Trp227/230 (e.g. AKR1D1 and AKR1C4),

in other cases one or more substitutions elsewhere are required to yield the same results. (e.g. AKR1C1 and AKR1C2). This seems to be further substantiated by the fact that simple substitution of loop residues directly contacting the substrate from one isoform to another was not sufficient to convert their specificity, but rather required loop chimeras (Ma and Penning 1999). This might be an indication that the ‘second shell’ of residues around those forming the substrate pocket would be limiting the side-chains to move and accommodate the ligand properly (i.e. induced fit). The loop chimeras might have remediated this limitation through an alternative path to reinstall the needed flexibility.

We were fortunate enough to have been able to validate our docking solutions against high-resolution crystallographic structures in complex with several steroids. Our protocol could reliably identify the correct poses and has placed the substrate ligand within 0.9Å of the equivalent atoms from cortisone (3CMF) and progesterone (3COT) (four rings and the keto group) (Fig. 3A, B). Furthermore, we believe that we have gathered a good body of evidence – from both docking predictions and analysis of experimental work – that Trp227/230 (switch tryptophan) is one of the most critical residues governing substrate specificity in several family 1 AKRs and that different substrate specificity and reaction can be achieved with small rearrangements in the side-chains rather than a major change in the loop conformation.

Finally, we would like to emphasise that *in silico* docking in conjunction with rational homology modelling and loop modelling can yield very reliable docking poses to help rationalise the varied binding modes presented by the members of this versatile family of enzymes. These hypothetical binding modes can then be used as guides in the design of experiments (e.g. site directed mutagenesis) that will allow validation of the model. We expect to be able to extend the use of modelling and docking to help interpret and analyse other members of the AKR family in order to understand the molecular mechanisms underlying the substrate specificity and reaction. This will ultimately provide the basis for prediction of new putative ligands and additional functions as well as to the design of new specificity and catalysis mechanisms for this diverse family of enzymes.

Our docking protocol could reliably identify the correct poses and has placed the substrate ligand within 0.9Å of the equivalent atoms from cortisone and progesterone (four rings and the keto group) (Fig. 3A,B). We would like to emphasise that in this case *in silico* docking in conjunction with rational homology modelling and loop modelling yielded very reliable docking poses to help rationalise the varied binding modes presented by the members of this versatile family of enzymes. Additionally, we believe that we have gathered a good body of evidence – from both docking predictions and analysis of experimental work – that Trp227/230 (switch tryptophan) is one of the most critical residues governing substrate specificity in several family 1 AKRs. Thus, we expect to be able to expand this knowledge to interpret and analyse other members of the AKR family in order to understand the molecular mechanism underlying the substrate specificity. This will ultimately provide the rationale for predicting new putative binders as well as additional functions of this diverse family of enzymes.

ACKNOWLEDGMENTS

We thank Trevor Penning for fruitful discussions. The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust.

REFERENCES

- Abagyan, R., M. Totrov, et al. (1994). "ICM: A New Method For Protein Modeling and Design: Applications To Docking and Structure Prediction From The Distorted Native Conformation." *J. Comp. Chem.* **15**: 488-506.
- Bauman, D. R., S. Steckelbroeck, et al. (2004). "The roles of aldo-keto reductases in steroid hormone action." *Drug News Perspect* **17**(9): 563-78.
- Cardozo, T., M. Totrov, et al. (1995). "Homology modeling by the ICM method." *Proteins* **23**(3): 403-14.
- Di Costanzo, L., J. E. Drury, et al. (2008). "Crystal Structure of Human Liver {Delta}4-3-Ketosteroid 5{beta}-Reductase (AKR1D1) and Implications for Substrate Binding and Catalysis." *J Biol Chem* **283**(24): 16830-9.
- Jez, J. M. and T. M. Penning (1998). "Engineering steroid 5 beta-reductase activity into rat liver 3 alpha-hydroxysteroid dehydrogenase." *Biochemistry* **37**(27): 9695-703.
- Jez, J. M. and T. M. Penning (2001). "The aldo-keto reductase (AKR) superfamily: an update." *Chem Biol Interact* **130-132**(1-3): 499-525.
- Jin, Y. and T. M. Penning (2006). "Molecular docking simulations of steroid substrates into human cytosolic hydroxysteroid dehydrogenases (AKR1C1 and AKR1C2): insights into positional and stereochemical preferences." *Steroids* **71**(5): 380-91.
- Jin, Y. and T. M. Penning (2007). "Aldo-keto reductases and bioactivation/detoxication." *Annu Rev Pharmacol Toxicol* **47**: 263-92.
- Ma, H. and T. M. Penning (1999). "Conversion of mammalian 3alpha-hydroxysteroid dehydrogenase to 20alpha-hydroxysteroid dehydrogenase using loop chimeras: changing specificity from androgens to progestins." *Proc Natl Acad Sci U S A* **96**(20): 11161-6.
- Oppermann, U., C. Filling, et al. (2003). "Short-chain dehydrogenases/reductases (SDR): the 2002 update." *Chem Biol Interact* **143-144**: 247-53.
- Palermo, M., M. G. Marazzi, et al. (2008). "Human Delta4-3-oxosteroid 5beta-reductase (AKR1D1) deficiency and steroid metabolism." *Steroids* **73**(4): 417-23.
- Penning, T. M. (1999). "Molecular determinants of steroid recognition and catalysis in aldo-keto reductases. Lessons from 3alpha-hydroxysteroid dehydrogenase." *J Steroid Biochem Mol Biol* **69**(1-6): 211-25.
- Russell, D. W. (2003). "The enzymes, regulation, and genetics of bile acid synthesis." *Annu Rev Biochem* **72**: 137-74.
- Steckelbroeck, S., Y. Jin, et al. (2004). "Human cytosolic 3alpha-hydroxysteroid dehydrogenases of the aldo-keto reductase superfamily display significant 3beta-hydroxysteroid dehydrogenase activity: implications for steroid hormone metabolism and action." *J Biol Chem* **279**(11): 10784-95.
- Totrov, M. and R. Abagyan (1997). "Flexible protein-ligand docking by global energy optimization in internal coordinates." *Proteins Suppl* **1**: 215-20.

Figure Legends

Figure 1: (A-D) The best poses achieved by the ligands in the active site after docking experiments. In each figure, the active site is shown in two different orientations – left: frontal view, right: side view (rotated 90 degrees along Y-axis). C4 of the nicotinamide ring is highlighted as a magenta wireframe sphere. **A:** AKR1D1 (model) + substrate; **B:** AKR1D1 (model) + product; **C:** AKR1C4 (modelled loop) + substrate; **D:** AKR1C4 (modelled loop) + product. **(E)** An example of non-productive poses (thin grey lines) found when models from the “set-1IHI” were used as receptors in the docking experiments. The model of AKR1D1 is shown. A catalytically active pose of the substrate is shown as an orange wireframe for comparison. The upper image shows the frontal view of the active site, the lower image shows the top view (rotated 45 degrees around X-axis) to illustrate how the ligands explore the pocket formed by loop B. **(F)** The pocket formed by loop B has different volumes depending on its residue composition. Green: AKR1C2 (1IHI) has a histidine in position 222 that protrudes inside the pocket, resulting in the restriction of the range of rotamer conformations adopted by the switch tryptophan. (Trp227/230). Magenta and cyan: respectively rat AKR1C9 (1AFS) and AKR1D1 (3BUR) have a serine in the same position, hence larger pockets, allowing the switch tryptophan to adopt an ‘inwards’ rotamer. The pink wireframe depicts the pose adopted by testosterone in complex with AKR1D1 (3BUR), where the larger pocket is used to accommodate the steroid.

Figure 2: Switch tryptophan rotamers: **(A)** Cluster of ‘across’ rotamers: AKR1C2 (1IHI – yellow, 1J96 – grey), AKR1C3 (1RY0 – orange), AKR1C4 (2FVL – red), AKR1D1 (3BV7 – lime, 3BUR – cyan); **(B)** Cluster of ‘inward’ rotamers: rat AKR1C9 (1AFS – cyan), model of AKR1D1 (set-1AFS – orange), AKR1D1 (3COT – purple, 3CMF – yellow); **(C)** AKR1C1 (1MRQ - green) – an example of an ‘across’ rotamer where the steroid has ring D close to the cofactor whilst presenting its beta side to the indole group of the “switch tryptophan”.

Figure 3: Comparison of structures of AKR1D1 + steroids with the docking solutions using homology models. Orange: homology model of AKR1D1 + 4-cholesten-7 α -ol-3-one (substrate); Yellow: AKR1D1 (3CMF) co-crystallised with cortisone; Purple: AKR1D1 (3COT) co-crystallised with progesterone. The C4 of the nicotinamide ring is highlighted as a magenta wireframe sphere. **(A)** Comparison of steroid poses after superimposition of the enzymes' backbone; **(B)** same as **(A)**, rotated 90 degrees around the Y-axis. **(C)** Overall fold comparison.

Table 1: Data collection and refinement statistics

PDB ID	2FVL
Space group	$P4_32_12$
Unit Cell	a = 166.00 Å, b=166.00 Å, c=194.94 Å
No. unique reflections	104763
Resolution (Å)	32.7-2.4 (2.46-2.4)
Completeness (%)	98.3
Redundancy	6.9 (4.6)
I/sigma	15.3 (2.0)
R _{meas}	13.7 (82.0)
R _{cryst} /R _{free}	0.166/0.205
RMSD bond length (Å)	0.011
RMSD bond angle (°)	1.42
No. of protein atoms	7816
No. of hetero atoms, ions	144
No. of water molecules	834

Highest resolution shell indicated in parenthesis

Table 2: Correlation between steroid binding poses and rotamer orientation adopted by the “switch tryptophan”. An example of each combination of receptor and ligand conformation is shown, along with the PDB code. Modeled poses are indicated.

	Orthogonal		Parallel
	Ring A in (3 α or 3-keto)	Ring D in (17 β /20 α or 17-/20-keto)	Rings long axis parallel to cofactor
Beta side up	<p>Across Switch Trp</p> <p>AKR1C2: testosterone (1J96)*</p>	<p>Across Switch Trp</p> <p>AKR1C2: UCDA (1IHI)*</p>	<p>Across Switch Trp</p> <p>AKR1D1: testosterone (3BUR)</p>
Alpha side up	<p>Inwards Switch Trp</p> <p>AKR1D1: cortisone (3CMF) AKR1D1: progesterone (3COT) AKR1C9: testosterone (1AFS) AKR1D1: 4-cholesten-7α-ol-3-one (model) AKR1C4: 5β-cholestan-7α-ol-3-one (model) AKR1C4: 5β-cholestan-3α,7α-diol (model)</p>	<p>Across Switch Trp</p> <p>AKR1C1: progesterone (1MRQ)*</p>	<p>Not yet seen</p>

* denotes the presence of a bulky histidine at the loop B pocket (see fig. 1F)

Steroid molecule (cortisone) for illustrative purpose only