FEBS Letters

FEBS PRESS science publishing by scientists

COMMUNICATION

Tolrestat acts atypically as a competitive inhibitor of the thermostable aldo-keto reductase Tm1743 from *Thermotoga maritima*

Chenyun Zhang[1], Zhenzhen Min[1], Xuemeng Liu[1], Chao Wang[1], Zhiguo Wang[1], Jiejie Shen[1], Wanrong Tang[1], Xin Zhang[1], Dan Liu[1] and Xiaoling Xu[1,2]

1 School of Medicine, Hangzhou Normal University, China
2 Institute of Cardiovascular Disease Research, The Affiliated Hospital of Hangzhou Normal University, China

Correspondence
X. Xu, School of Medicine, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China
Fax/Tel: +86 571 28861723
E-mail: xuxl@hznu.edu.cn

Chenyun Zhang, Zhenzhen Min, Xuemeng Liu, and Chao Wang contribute equally to this work

(Received 29 July 2019, revised 17 September 2019, accepted 29 September 2019, available online 17 October 2019)

doi:10.1002/1873-3468.13630

Edited by Judit Ovádi

Tolrestat and epalrestat have been characterized as noncompetitive inhibitors of aldo-ketone reductase 1B1 (AKR1B1), a leading drug target for the treatment of type 2 diabetes complications. However, clinical applications are limited for most AKR1B1 inhibitors due to adverse effects of cross-inhibition with other AKRs. Here, we report an atypical competitive binding and inhibitory effect of tolrestat on the thermostable AKR Tm1743 from *Thermotoga maritima*. Analysis of the Tm1743 crystal structure in complex with tolrestat alone and epalrestat-NADP$^+$ shows that tolrestat, but not epalrestat, binding triggers dramatic conformational changes in the anionic site and cofactor binding pocket that prevents accommodation of NADP$^+$. Enzymatic and molecular dynamics simulation analyses further confirm tolrestat as a competitive inhibitor of Tm1743.

Keywords: aldo-ketone reductase; competitive inhibitor; diabetes complications; epalrestat; tolrestat

The aldo-ketone reductase (AKR) superfamily of NADPH-dependent oxidases is primarily responsible for catalyzing the reversible reduction of carbonyl groups in a wide range of aldehydes and ketones to their corresponding alcohols [1,2]. There are more than 190 known AKR enzymes that fall into 16 families and which are widely distributed across numerous metabolic processes, such as xylose metabolism [3], vitamin C biosynthesis [4], polyketide biosynthesis [5], and steroid metabolism [6]. Human AKR enzymes play central roles in the bioactivation or detoxification of drugs, carcinogens, and reactive aldehydes [7,8]; for this reason, they are categorized as phase I drug-metabolizing enzymes [1]. Several human AKRs, such as AKR1B1, AKR1C1-1C3, AKR1D1, and AKR1B10, have been implicated in diabetes complications [9–12], steroid hormone-dependent malignancies [13–15], bile acid deficiency [16,17], and defects in retinoic acid signaling [18–20]. As such, extensive exploration has been conducted to identify targeted inhibitors of AKRs that may help in curing these diseases.

A structurally diverse set of compounds has been found to inhibit the enzymatic activity of several disease-related AKRs. Among these, tolrestat [21] and epalrestat [22] are two powerful acetic acid inhibitors of AKR1B1 (aldose reductase), a leading drug target for type 2 diabetes complications, due to its ability to convert high circulating concentrations of glucose into sorbitol in non-insulin-dependent tissues [23,24]. Tolrestat also inhibited *in vivo* and *in vitro* activities of AKR1B10 [19], which is used as a marker for tumorigenesis and is essential in regulating the conversion of

Abbreviations
AKR, aldo-ketone reductase; AU, asymmetric unit; AUC, analytical ultracentrifugation; MD simulation, molecular dynamics simulation; NMA, normal mode analysis; Tm, *Thermotoga maritima*.

all-*trans*-retinaldehyde to retinol [18,20,25]. However, development of effective AKR1B inhibitors for clinical applications has been limited by cross-inhibition with other AKRs, caused by the highly conserved inhibitor-binding mode [26,27]. Thus, it is necessary to have a clear understanding of the inhibitor–enzyme interactions when designing inhibitors for a specific AKR.

Tolrestat and epalrestat have long been characterized, enzymatically and structurally, as noncompetitive inhibitors. Both inhibitors bind to AKR1B1 and AKR1B10 in complex with the enzyme cofactor $NADP^+$, which is a common feature shared by several inhibitor-bound AKR structures [19,28–31]. In these structures, the hydrophobic ring of tolrestat and epalrestat is stacked by aromatic residue side chains located in a so-called 'specificity pocket'. The hydrophilic tails of the inhibitors are typically anchored to an 'anionic site' composed of strictly conserved Tyr, His residues, and the $NADP^+$ nicotinamide moiety that interact with the carboxyl group through hydrogen bonding. The $NADP^+$ cofactor is bound in an extended *anti*-conformation in a conserved binding pocket adjacent to the active sites that contains an Asp-Lys-His-Tyr catalytic tetrad.

In previously published work, we characterized the activity and crystal structure of the novel AKR Tm1743, cloned from *Thermotoga maritima,* in complex with $NADP^+$ at 2.0 Å resolution. We investigated its enantioselectivity in catalyzing the asymmetric synthesis of the angiotensin drug intermediate ethyl (R)-2-hydroxy-4-phenylbutyrate through semirational enzyme design [32]. We found that the Trp21 and Trp86 residues located in the specificity pocket play crucial roles in determining the enantioselectivity of Tm1743. As a native AKR, Tm1743 exhibits some essential properties of an ideal biocatalyst, including high thermostability, strong chemical tolerance, and activity toward a broad range of ketone and aldehyde substrates [33]. These findings therefore warranted further structural and functional characterization of the interactions between Tm1743 and the inhibitors used in diabetes treatment for potential development in pharmaceutical applications.

Here, we report the crystal structure of Tm1743 in complex with the inhibitors tolrestat at 1.6 Å and epalrestat-$NADP^+$ at 1.9 and 2.07 Å resolution, respectively. Compared to mammalian AKRs, Tm1743 immobilizes tolrestat alone in an atypical binding mode that triggers conformational changes of the anionic site and cofactor binding pocket. As a result, the hydrophilic tail of tolrestat is repositioned to occupy the nicotinamide binding region of the anionic site. Concomitantly, the cleft that accommodates the adenosine ring of $NADP^+$ is closed due to conformational changes of the amino acids that maintain the cleft structure. In contrast, epalrestat binding does not induce conformational changes that preclude the interaction with $NADP^+$. The respective binding modes of tolrestat and epalrestat were further confirmed by enzymatic reduction of ethyl acetoacetate by Tm1743 and molecular dynamics (MD) simulation analyses of the complex structures. This work thus revealed a previously unrecognized binding mode and inhibitory effect of tolrestat on an AKR enzyme. Our work provides the structural basis for the competitive inhibition of tolrestat on AKR enzymes. It will expand the diversity of inhibition mechanisms of AKR inhibitors and also contribute to their future pharmaceutical applications.

Materials and methods

Protein expression and purification

The gene sequence encoding Tm1743 from *T. maritima* was inserted into pET28a vector between the *NdeI* and *HindIII* restriction sites [33]. A plasmid with the correct insert confirmed by sequencing was then transformed into *Escherichia coli* BL21 (DE3) cells for expression of the N-terminal His_6-tagged recombinant protein. The transformed cells were grown in 1 L of LB medium containing 100 $mg{\cdot}mL^{-1}$ kanamycin at 37 °C until the OD_{600} reached 0.6–0.8. The gene expression was then induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) overnight at 25 °C.

Cells were harvested by centrifugation at 7903 ***g*** for 15 min at 4 °C and resuspended in 40 mL wash buffer containing 25 mM Tris/HCl pH 8.0, 150 mM NaCl, 20 mM imidazole, and 0.1 mM phenylmethylsulfonylfluoride prior to homogenization with a high-pressure homogenizer (Union, People's Republic of China). The insoluble cell debris was removed by centrifugation at 23 269 ***g*** for 40 min at 4 °C. The supernatant containing crude soluble proteins was boiled at 100 °C for 10 min to precipitate the majority of *E. coli* proteins. After centrifugation at 34 310 ***g*** for 20 min, the supernatant containing primarily Tm1743 protein was loaded onto a Ni^{2+}-chelating affinity chromatography column (2 mL; GE Healthcare, Waukesha, WI, USA) and was rinsed with 100 mL buffer A (25 mM Tris–HCl pH 8.5, 200 mM NaCl, 20–40 mM imidazole) to remove nonspecifically bound proteins. The bound Tm1743 protein was eluted with buffer B (25 mM Tris–HCl pH 8.5, 200 mM NaCl, 50–100 mM imidazole). The eluates were further purified by a HiLoad 16/600 Superdex 200 PG size exclusion column (GE Healthcare) to 95% purity.

Crystallization of Tm1743 in complex with tolrestat and epalrestat

The purified Tm1743 enzyme was concentrated to 20 mg·mL^{-1} in buffer (25 mM Hepes pH 7.5 and 50 mM NaCl) at 4 °C using an Amicon Ultra centrifugal filter device (10-kDa molecular weight cutoff; Millipore, Danvers, MA, USA). Protein concentration was determined using a NanoDrop (Thermo Scientific, Waltham, MA, USA) to measure absorption at 280 nm. The enzyme was incubated with $NADP^+$ and individual inhibitors (tolrestat or epalrestat) at a 1 : 1.5 : 1.5 molar ratio for 2 h before crystallization. Crystallization was performed using the hanging-drop vapor diffusion method at 4 °C. 1.5 μL of protein sample was mixed with an equal volume of the reservoir solution, and the mixture was equilibrated against 200 μL reservoir solution. Crystals of tolrestat-bound Tm1743 were obtained at 15 mg·mL^{-1} with the reservoir solution (10% PEG10 000, 0.1 M sodium acetate trihydrate pH 4.5) at 16 °C. The epalrestat-bound Tm1743 was crystallized with 10% Jeffamine M600 pH 7.0 and 0.1 M sodium acetate trihydrate pH 4.0 at 16 °C. The crystals of epalrestat-bound Tm1743 in $P3_221$ space group were obtained in reservoir solution containing 10% PEG 20 000, 0.2 M $MgCl_2$, and 0.1 M sodium acetate trihydrate pH 4.5 at 16 °C.

Crystal data collection, structure determination, and refinement

The optimized Tm1743 crystals were cryoprotected by adding 25% glycerol to the reservoir solution and flash-freezing with liquid nitrogen. A 1.6 Å resolution dataset of tolrestat-bound Tm1743 was collected at SSRF beamline BL18U1 with a 300-mm crystal-to-detector distance at wavelength of 0.97861 Å. Three hundred and sixty diffraction frames were collected with 1° oscillation per image. The crystal belongs to space group $P3_121$ with unit cell dimensions $a = b = 84.594$ Å, $c = 92.990$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The crystal diffraction data of epalrestat-bound Tm1743 were also collected at SSRF BL18U1 under the same conditions. A 1.9 Å dataset was collected belonging to space group $P3_121$ with unit cell dimensions $a = b = 84.172$ Å, $c = 93.166$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The diffraction dataset of the crystal grown in 10% PEG 20 000 was collected at 2.07 Å resolution, with space group $P3_221$ and unit cell dimensions $a = b = 84.128$ Å, $c = 185.765$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$ (Table 1).

Diffraction data were processed, integrated, and scaled with HKL3000R software [34]. The data quality was assessed using SFCHECK [35], and the solvent content was calculated using MATTHEWS_COEF from CCP4 [36,37]. The inhibitor-bound structures were determined by the molecular replacement method using the coordinates of $NADP^+$-bound Tm1743 (PDB ID 5DAN) [32] as a search model. The PHASER program [38] from the CCP4 package was employed to determine the initial phases; iterative model building and refinement were performed using COOT [39] and REFMAC5 [40] to obtain the final model (Table 1).

Enzymatic assays

The enzymatic activity of Tm1743 in reducing the soluble ethyl acetoacetate substrate was assayed spectrophotometrically at 30 °C by measuring the change in the absorbance of NADPH at 340 nm ($\varepsilon = 6.22$ mM^{-1}·cm^{-1}). The standard assay mixture (200 μL) was composed of 50.0 mM Tris/HCl pH 9.0, 0–25.0 mM ethyl acetoacetate, 0.12 mM NADPH, and 4.0 μM Tm1743 enzyme. One unit of enzymatic activity was defined as the amount of enzyme that catalyzes the oxidation of 1.0 μmol NADPH per min. The apparent Michaelis–Menten constants (K_m) and V_m were measured by Lineweaver–Burk double-reciprocal plots of the reaction velocity and substrate concentrations at 0, 1.5, 2.5, 5.0, 10.0 and 25.0 mM (Table 2).

The inhibitory effects of tolrestat and epalrestat were investigated by double-reciprocal plot of the reaction velocity and substrate concentrations, by varying the concentrations of tolrestat at 0, 20.0, 40.0, 50.0, and 60.0 nM, and the concentrations of epalrestat at 0, 25.0, 50.0, 60.0, and 75.0 nM. The values of K_m and V_m at different inhibitor concentrations were the apparent K_m and V_m, hereinafter denoted as K_m^{app} and V_m^{app}, respectively. The apparent inhibitory constant (K_i) of tolrestat and epalrestat was calculated by plotting the K_m^{app}/V_m^{app} against different inhibitor concentrations. To check whether NADPH could displace the bound tolrestat from Tm1743 enzyme, the enzymatic assay was carried out in reaction mixture (200 μL) with 50.0 mM Tris/HCl pH 9.0, 4.0 μM Tm1743, 25.0 mM ethyl acetoacetate, 25 nM tolrestat, and NADPH concentrations at 50, 100, and 150 nM.

Molecular dynamics simulation and binding free energy calculations

Molecular dynamics simulations were performed by using the AMBER 12 software [41]. The tolrestat-, $NADP^+$- and epalrestat-$NADP^+$-bound Tm1743 structures were individually immersed into the center of a truncated octahedron box of TIP3P water molecules with a margin distance of 12.0 Å. The environmental sodium counterions were added to keep the system in electric neutrality. The AMBER ff14SB force field was applied for Tm1743 [42]. The force field parameter of $NADP^+$ was retrieved from previous report [43]. For the inhibitor tolrestat and epalrestat, their atomic partial charges were calculated using the restricted electrostatic potential (RESP) method with a basis set of HF/6-31G(d) on the DFT B3LYP/6-31G(d) optimized structures [44]. By following the same procedure in our previous report [45], each simulation was conducted with a time scale of 60 ns.

Upon the equilibrium of MD simulations, the binding free energies (ΔG_{bind}) between Tm1743 and bound

Table 1. Data collection and refinement statistics.

	Tm 1743 bound with tolrestat (6KIK)	Tm 1743 bound with epalrestat and $NADP^+$_1 (6KIY)	Tm 1743 bound with epalrestat and $NADP^+$_2 (6KY6)
Data collection			
Diffraction source	BL18U1, SSRF	BL18U1, SSRF	BL18U1, SSRF
Wavelength (Å)	0.97861	0.97861	0.97861
Space group	$P3_121$	$P3_121$	$P3_221$
Cell parameters (Å)	$a = b = 84.594$ Å $c = 92.990$ Å $\alpha = \beta = 90°$ $\gamma = 120°$	$a = b = 84.172$ Å $c = 93.166$ Å $\alpha = \beta = 90°$ $\gamma = 120°$	$a = b = 84.128$ Å $c = 185.765$ Å $\alpha = \beta = 90°$ $\gamma = 120°$
Resolution (Å)	50 (1.63)[a] – 1.60	50 (1.93) – 1.90	50 (2.11) – 2.07
Total no. of reflections	1 004 056 (50 700)	588 966 (27 844)	492 392 (24 440)
No. of unique reflections	51 163 (2535)	30 676 (1497)	47 317 (2350)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy	19.6 (20.0)	19.2 (18.6)	10.4 (10.4)
$I/\sigma(I)\rangle$	5.3 (5.5)	27.2 (4.9)	26.0 (5.4)
R_{merge} (%)	10.5 (43.6)	16.4 (75.0)	17.7 (92.8)
Overall *B* factors from Wilson plot ($Å^2$)	19.0	14.8	19.6
Matthews coefficient	3.04	3.02	3.12
V_M ($Å^3 \cdot Da^{-1}$)			
Solvent content (%)	59.6	59.3	60.6
Refinement			
Resolution (Å)	50.0–1.60	50.0–1.90	50.0–2.07
Total no. of reflections	51 118	29 156	44 894
No. of reflections used	48 525	27 668	42 530
R_{work}/R_{free} (%)[b]	17.3/19.1	17.2/20.3	21.7/25.6
No. of atoms	2585	2365	4857
Protein	2212	2212	4426
Tolrestat	1	0	0
Epalrestat	0	1	2
$NADP^+$	0	1	2
Cl^-	0	0	5
Water	332	92	305
Wilson B factors ($Å^2$)	16.8	18.3	22.4
Average B factors ($Å^2$)	22.0	21.0	28.7
Protein	20.2	20.9	27.5
Ligands	21.5	32.8	44.8
R.m.s. deviations			
Bond angle (°)	0.83	1.40	1.46
Bond length	0.005	0.009	0.009
Ramachandran plot			
Favored (%)	98.9	99.6	99.3
Allowed (%)	1.1	0.4	0.7
Outliers (%)	0.0	0	0

[a] Values in parentheses are for the highest resolution shell.; [b] R_{free} was calculated using 5% of randomly selected reflections that were excluded from the refinement.

inhibitors tolrestat, epalrestat, and cofactor $NADP^+$ were obtained through the molecular mechanics/generalized Born surface area (MM/GBSA) calculation approach [46]:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \quad (1)$$

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (2)$$

$$\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{GB} + \Delta G_{SA} \quad (4)$$

where E_{MM} is the gas phase interaction energy comprising van der Waals energy (E_{vdW}) and electrostatic energy (E_{ele}). G_{solv} is the solvation free energy,

Table 2. Kinetic parameters of the Tm1743 alone and with the addition of inhibitors tolrestat and epalrestat in reducing ethyl acetoacetate.

Enzymes	V_m (μM·s^{-1})	K_m (mM)	*kcat* (s^{-1})
Tm1743	0.87 ± 0.02	14.46 ± 0.50	0.22 ± 0.00
Tolrestat (nM)[a]			
20	0.90 ± 0.02	18.72 ± 0.91	0.23 ± 0.01
40	0.94 ± 0.04	22.20 ± 1.52	0.24 ± 0.02
50	0.94 ± 0.05	24.75 ± 2.23	0.23 ± 0.01
60	0.97 ± 0.03	27.51 ± 1.41	0.24 ± 0.01
Epalrestat (nM)[a]			
25	0.91 ± 0.04	28.61 ± 2.17	0.23 ± 0.01
50	0.92 ± 0.04	37.63 ± 2.45	0.23 ± 0.01
60	0.86 ± 0.06	40.79 ± 4.33	0.22 ± 0.01
75	0.82 ± 0.08	44.98 ± 6.19	0.21 ± 0.02

[a] The values of K_m and V_m at different inhibitor concentrations are the apparent K_m and V_m, hereinafter denoted as K_m^{app} and V_m^{app}, respectively.

including the contributions form a polar part (G_{GB}) and a nonpolar part (G_{SA}). ΔG_{GB} was estimated using the generalized Born model with the interior and exterior dielectric constants set to 4 and 80, respectively [47]. ΔG_{SA} was estimated using the LCPO algorithm: $\Delta G_{SA} = \gamma \Delta SASA + \beta$, where γ and β were set to 0.0072 and 0, respectively [48]. By performing the normal mode analysis (NMA), $T\Delta S$ that represents the entropy contribution was estimated using the NMODE module. Two hundred snapshots were evenly extracted from the last 20 ns trajectories for the calculations of ΔE_{vdW}, ΔE_{ele}, ΔG_{GB}, and ΔG_{SA}. For the calculation of entropy, only 40 snapshots were evenly extracted from the last 20 ns trajectories due to the expensive computational cost of NMA [49].

Gel filtration and sedimentation velocity analytical ultracentrifugation

To check the oligomerization state of Tm1743 in the solution, we performed gel filtration and sedimentation velocity analytical ultracentrifugation (AUC) of Tm1743 alone, and in complex with inhibitors and cofactor NADP$^+$, respectively. The Tm1743 enzyme (23 mM) was incubated with NADP$^+$ and individual inhibitors (tolrestat or epalrestat) at a 1 : 1.5 : 1.5 molar ratio for 2 h before gel filtration and AUC analyses. Then, the samples were loaded on a Superdex 75 10/300 GL size exclusion column (GE Healthcare) and eluted with buffer containing 25 mM Tris/Cl pH 8.5, 20 mM NaCl at 0.4 mL·min^{-1}. Sedimentation experiments were performed on a Beckman Coulter ProteomeLab XL-I ultracentrifuge using a 4-hole An-60Ti rotor. Samples with an initial absorbance at 280 nm of approximately 0.5–0.8 were equilibrated for 2 h at 20 °C under a vacuum prior to sedimentation. The absorbance at 280 nm was measured using a continuous scan mode during sedimentation at 220 220 ***g*** in 12 mm double-sector cells. The data were analyzed using sedfit.

Results

Tm1743 inhibitor-binding mode is different for tolrestat than epalrestat and NADP$^+$

The crystal structure of tolrestat-bound Tm1743 (PDB 6KIK) was determined by the molecular replacement method and refined to an R_{work} of 17.3% and R_{free} of 19.1% at 1.6 Å resolution (Table 1). The *P3$_1$21* trigonal crystal contains only one Tm1743 molecule per asymmetric unit (AU). The refined structure covers full-length Tm1743 (Met1-Gly274) which is composed of an N-terminal β hairpin (β1 and β2), a central β_7/α_8 TIM barrel, and two additional α helices, α8 (Glu208-Gln220) and α11 (Glu264-Ser272), in the C-terminal domain (Fig. 1A). Unlike typical AKR enzymes, Tm1743 lacks one β strand in the TIM barrel, and thus, its C terminus possesses a stable α-helix (α11) instead of long, random coils.

To obtain the complex structure of tolrestat-bound Tm1743, we incubated Tm1743 with the enzyme cofactor NADP$^+$ and tolrestat before crystallization. However, no electron density indicating the presence of NADP$^+$ was observed in the cofactor binding site, and instead, a tolrestat molecule was resolved in the loop regions above the TIM barrel (Fig. 1A and Fig. S1A). The tolrestat molecule entered the active-site pocket of Tm1743 from the upper face of the TIM barrel, with its polar tail inserted into the anionic site and the hydrophobic head pointing outwards the TIM barrel (Fig. 1B). The hydrophobic ring of tolrestat was stacked against the indole rings of Trp86 and Trp21; its trifluoromethyl group was located within a 3.0 Å distance of the hydroxyl group of Glu27 at the TIM barrel surface. The carboxyl group was directed into a hydrophilic pocket formed by Tyr57 and the catalytic tetrad Asp53-Lys84-His117-Tyr58, wherein it was hydrogen-bonded with the tyrosyl group of Tyr57 (2.7 Å) and the amine group of Lys84 (3.0 Å) (Fig. 1B).

The crystal structure of Tm1743 in complex with NADP$^+$ and epalrestat (PDB 6KIY) was determined and refined to an R_{work} of 17.2% and R_{free} of 20.3% at 1.9 Å resolution (Table 1). Another crystal structure with space group of *P3$_2$21* (PDB 6KY6) was refined to R_{work} of 21.7% and R_{free} of 25.6% at 2.07 Å resolution (Table 1). The *P3$_2$21* crystal contains two Tm1743 monomers per AU, each monomer bound with one epalrestat and one NADP$^+$ molecules that symmetrically sit on the two sides of the two fold axis (Fig. S2).

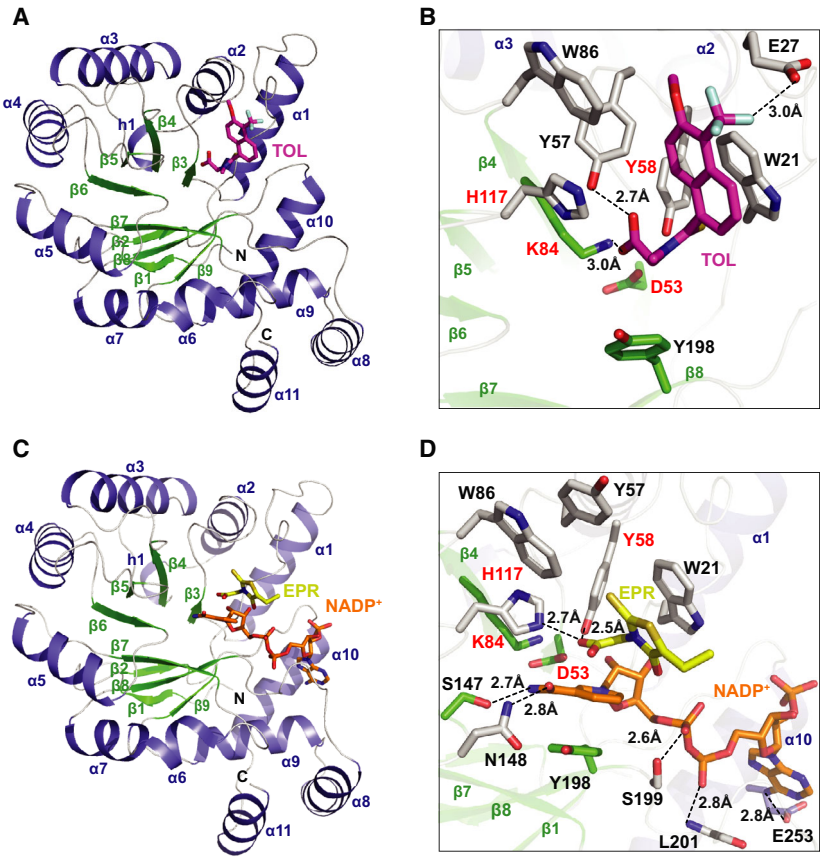

Fig. 1. Tm1743 inhibitor-binding mode is different for tolrestat than epalrestat and NADP$^+$. (A) Overall structure of tolrestat-bound Tm1743 (PDB 6KIK). Tolrestat (magenta stick model) enters the Tm1743 active-site pocket from the upper face of the TIM barrel, with its polar tail inserted into the barrel and the hydrophobic head pointing outwards. The secondary structures are colored in blue (α helices), green (β strands), and gray (coils). (B) Binding of tolrestat in the loop regions of the TIM barrel. The hydrophobic ring is stacked against the indole rings of Trp86 and Trp21; the hydrophilic tail is hydrogen-bonded with the Tyr57 and Lys84 side chains. The Asp53-Lys84-His117-Tyr58 catalytic tetrad (red) and amino acid residues essential for coordinating tolrestat are shown as stick models; hydrogen bonding is indicated by a black dashed line. (C) Overall structure of Tm1743 bound to epalrestat in complex with NADP$^+$ (PDB 6KIY). The epalrestat (yellow) is immobilized above the TIM barrel that harbors the NADP$^+$ (orange) binding pocket. (D) Immobilization of epalrestat and NADP$^+$ in Tm1743. The hydrophobic ring of epalrestat (yellow) is trapped in the specificity pocket formed by Trp21, Trp86, and Tyr57; the carboxyl group is hydrogen-bonded with His117 and Tyr58. NADP$^+$ (orange) is coordinated by extensive hydrogen bonding interactions in the cofactor binding pocket. The amino acid residues essential for coordinating epalrestat and NADP$^+$ are shown as stick models; dashed lines represent hydrogen bonding interactions.

Gel filtration and AUC analyses demonstrated that either Tm1743 alone or in complex with tolrestat, epalrestat, and NADP$^+$ existed as a monomer in the solution (Fig. S3). Therefore, the monomer association observed in the *P3$_2$21* crystal of epalrestat-bound Tm1743 is a result of crystal packing.

In the two epalrestat-bound structures, both the NADP$^+$ cofactor and the epalrestat head were resolved with clear electron densities (Fig. 1C and Figs S1B and S2C). The hydrophobic ring of epalrestat was trapped in the specificity pocket formed by side chains of Trp21, Trp86, and Tyr57; it was parallel with the indole rings of Trp21 and Trp86 and perpendicular to the tryptophanyl ring of Trp118. The carboxyl group of epalrestat was immobilized by hydrogen bonding interactions with His117 (2.7 Å) and Tyr58 (2.5 Å) in the anionic site, where the nicotinamide ring of NADP$^+$ was stacked against the side chain of Tyr198, leaving the other side of the nicotinamide ring facing the catalytic tetrad from above (Fig. 1D). NADP$^+$ was coordinated into the cofactor binding pocket through extensive hydrogen bonding interactions with residues from β6, β8, β9, and α10 (Fig. 1D).

Tm1743 contains an atypical tolrestat binding pocket

Both tolrestat and epalrestat have been previously characterized as noncompetitive AKR inhibitors that were bound in complex with the enzyme cofactor $NADP^+$ [19,28–31]. To investigate the mechanisms of tolrestat binding with Tm1743, we compared all of the deposited tolrestat-bound AKR structures. Superimposition analyses showed that these AKR structures matched well at the central TIM barrel, as well as with the conformations of bound $NADP^+$ molecules. Interestingly, the tolrestat molecule was immobilized in a completely distinct conformation in Tm1743 compared to its conformation in mammalian AKRs; the hydrophobic ring was rotated about 60° relative to its position in mammalian AKRs (Fig. 2A).

All mammalian AKRs examined in this study contained three additional loops not found in Tm1743 which surround the tolrestat binding sites, namely loop A, loop B, and loop C. Loop C is comprised of the 24-amino acid extension found at the C terminus of mammalian AKRs that fold into one or two short α-helices (Fig. 2A, Figs S4A and S5). The extended loop A and loop C were essential for accommodating the tolrestat ring in a narrow specificity pocket, while loop B was involved in binding the pyrophosphate group and adenine ring of $NADP^+$ (Fig. 2A, Fig. S6A,B).

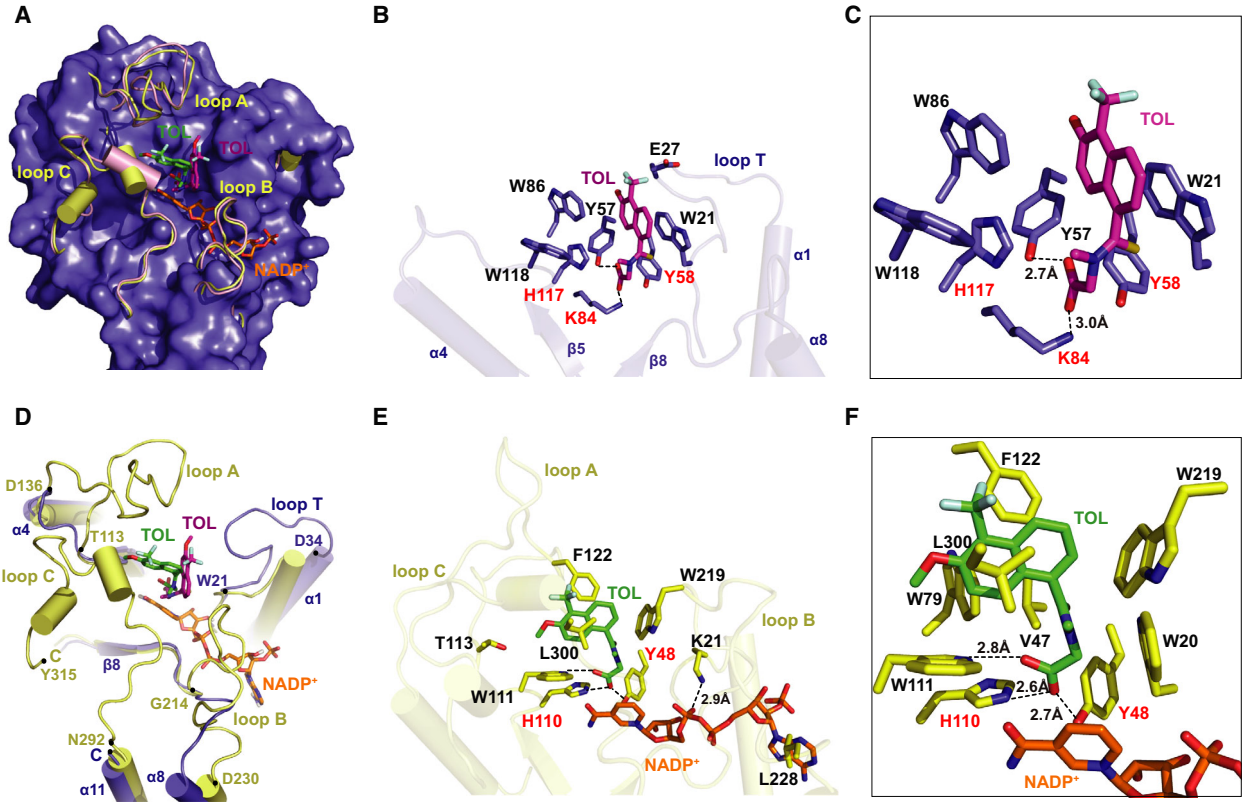

Fig. 2. The tolrestat binding pocket of Tm1743 and mammalian AKR1B1. (A) Superimposition of Tm1743 (blue, PDB 6KIK) with AKR1B1 (yellow, PDB 2FZD) and AKR1B10 (pink, PDB 1ZUA). The space-filling model of the Tm1743 structure is shown with the three additional loops (loops A, B, and C) present in mammalian AKRs. Bound tolrestat is shown as a stick model in Tm1743 (magenta) and in mammalian AKRs (green). (B) The tolrestat binding pocket in Tm1743. The tolrestat molecule (magenta) is immobilized in a pocket formed by loop T and residues located in the loops α4-β5 and β8-α8. The hydrophobic ring of tolrestat is stacked by Trp21 and Trp86 in the specificity pocket, and the hydrophilic tail was hydrogen-bonded with Tyr57 and Lys84 in the anionic site. The critical secondary structures (ribbon diagram) and specific amino acids (stick models) that form the binding pocket are shown with tolrestat and enlarged in (C). (D) Superimposition of tolrestat-bound Tm1743 (blue) with AKR1B1 (yellow). The three additional loops in mammalian AKRs are positioned around the tolrestat binding pocket, while loop T covers the tolrestat binding pocket from the top in Tm1743. Residues located at the terminus of a loop are shown as black dots. (E) Immobilization of tolrestat (green) and $NADP^+$ (orange) in human AKR1B1. Three residues (Phe122, Trp219, and L300) bring the flipped loops A, B, and C together to cover the top of the tolrestat binding pocket. The hydrophobic ring of tolrestat is deflected into the cleft between loop A and loop C, while the hydrophilic tail is hydrogen-bonded (dashed lines) with Trp111, His110, and Tyr48 in the anionic site. Lys21 and Leu228 are essential for binding the $NADP^+$ molecule. (F) Enlargement of tolrestat and cofactor interactions with binding pocket residues in human AKR1B1.

Additionally, an insertion (Gly24-Asp34), designated as loop T (Gly24-Asp34), was present only in Tm1743 and absent in mammalian AKRs (Fig. 2B,D, Figs S4A and S5). Due to absence of the three additional loops, loop T in Tm1743 was flipped toward the tolrestat binding pocket from above. Within this loop, Glu27 was positioned on the tip close to the trifluoromethyl group, and the indole ring of Trp21 was essential for stacking the tolrestat ring in a flat conformation through steric hindrance with Trp86 on the opposite side of the tolrestat hydrophobic ring (Fig. 2B,C). As a result, the hydrophilic tail moved down into the anionic site, where the side chains of Trp118 and His117 had been rotated away, and there the carboxylate was hydrogen-bonded by a nonconserved Tyr57 residue and the Lys84 catalytic residue (Fig. 2B,C).

In mammalian AKR1B1 (PDB 2FZD and 1AH3), the hydrophobic ring of tolrestat was immobilized in a hydrophobic pocket formed by Phe122 from loop A (Thr113-Asp136), Trp219 from loop B (Gly214-Asp230), and L300 from loop C (Asn292-Tyr315) (Fig. 2D). These three residues (Phe122, Trp219, and L300) were positioned in a triangle 8 Å apart, bringing the flipped loops A, B, and C together to cover the top of the tolrestat binding pocket (Fig. 2E). The steric hindrance generated by the phenyl ring of Phe122 and the indole ring of Trp219 deflected the tolrestat hydrophobic ring into the cleft between loop A and loop C. Appearance of these three extended loops lifted the tolrestat ring up above the anionic site, wherein the backbone of the hydrophilic tail was accommodated by the side chains of Trp79 (Trp86 in Tm1743) and Val47 (Tyr57 in Tm1743) on one side of the hydrophilic tail, and Trp20 (Trp21 in Tm1743) on the opposite side of the tail (Fig. 2F).

The fact that Val47 in mammalian AKR1B1 is positionally equivalent to Tyr57 in Tm1743 prevents the hydrogen bonding interaction necessary for immobilization of the tolrestat carboxyl group, which was hydrogen-bonded with Trp111 (2.8 Å), His110 (2.6 Å), and Tyr48 (2.7 Å) (Fig. 2F). In contrast, the side chains of Trp86, Trp21, Trp118, and His117 in tolrestat-bound Tm1743 all underwent obvious conformational changes as that in AKR1B1 (Fig. S4B). The tolrestat binding mode in AKR1B10 (PDB 1ZUA) thus resembles that of AKR1B1 (Fig. S4C–F).

The hydrophilic tail of tolrestat occupied the anionic site in Tm1743

In contrast to tolrestat, the inhibitor epalrestat bound to $NADP^+$-Tm1743 in a conformation that is different to human AKR1B1 (PDB 4JIR) and AKR1B10 (PDB 4JIH) (Fig. 3A,B, Fig. S7A–D). Although the hydrophobic ring adopted different conformations from mammalian AKR1B1 and AKR1B10, the hydrophilic tails of the epalrestat were all immobilized by the highly conserved Trp, His, and Tyr residues above the anionic site, leaving enough space to accommodate the nicotinamide ring of $NADP^+$ (Fig. 3C and Fig. S7). Conformational changes in the epalrestat ring in human AKR1B10 resulted from the planar flip of the Trp112 side chain, which determines the inhibitor specificity between AKR1B1 and AKR1B10 [31]. The catalytic tetrad residues Asp53, Lys84, Tyr58, and His117 did not undergo conformational changes upon epalrestat binding (Fig. 3C). However, the side chain orientations of Tyr57, His117, and Trp118 in tolrestat-bound Tm1743 are dramatically different from that in AKR1B1 and AKR1B10, as well as epalrestat-bound Tm1743. In particular, the side chain orientation of Trp118 in tolrestat-bound Tm1743 resembles that of Trp112 in epalrestat-bound AKR1B10 (Fig. 3D).

Superimposition of the Tm1743 inhibitor-binding pockets revealed that the orientations of the Tyr57, His117, and Trp118 side chains play functional roles in immobilizing tolrestat and epalrestat in different binding modes (Fig. 4A). The hydrophobic rings of tolrestat and epalrestat were both sandwiched between the indole rings of Trp21 and Trp86, which did not undergo dramatic conformational changes, whereas in the tolrestat-bound structure, the side chains of His117 and Trp118 were all rotated away to broaden the anionic site (Fig. 4A). The longer hydrophilic tail of tolrestat was then mobilized into the nicotinamide ring binding region, where it was stabilized by the drastically flipped tyrosyl group of Tyr57 and the catalytic residue Lys84 (Fig. 4A). Interestingly, these essential inhibitor-binding residues share the same side chain orientations in epalrestat- and $NADP^+$-bound (PDB 5DAN) structures but not in the tolrestat-bound Tm1743 (Fig. 4B).

Tolrestat binding dismantled the cofactor binding pocket in Tm1743

To further elucidate the structural basis underlying the absence of $NADP^+$ in tolrestat-bound Tm1743, we superimposed the structure with a previously determined $NADP^+$-bound structure (PDB 5DAN). The overall tolrestat- and $NADP^+$-bound structures matched well with a main chain r.m.s.d. of 0.261 Å. Minor structural differences were observed at β5 and the connecting loop between β8 and α8 (corresponding to the mammalian loop B; Fig. 5A). Upon tolrestat

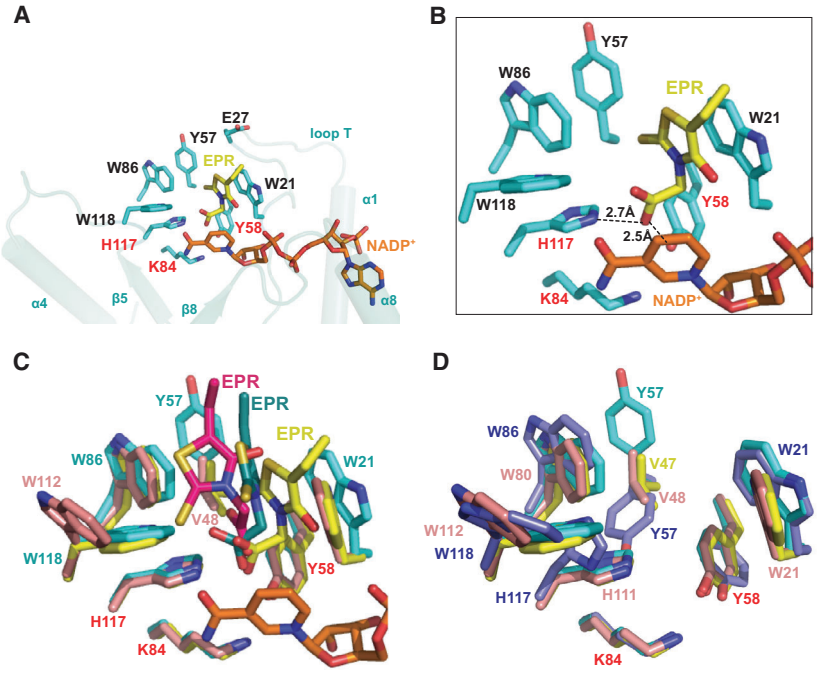

Fig. 3. Comparison of the epalrestat binding in Tm1743 with mammalian AKRs. (A) Overview of the epalrestat binding pocket in Tm1743 (PDB 6KIY). The epalrestat molecule (yellow) is immobilized in a pocket formed by loop T and residues located in loops α4-β5 and β8-α8. $NADP^+$ (orange) is bound in the cofactor binding pocket located in the TIM barrel. The secondary structures (cyan ribbons) and amino acids (cyan stick models) that participate in forming the epalrestat binding pocket are shown. (B) Enlargement of the epalrestat (yellow) and cofactor (orange) interactions with the binding residues in Tm1743. The hydrophobic ring of epalrestat is immobilized by Trp21, Tyr57, and Trp86 (cyan) in the specificity pocket, and the hydrophilic tail is hydrogen-bonded (dashed lines) with His117 and Tyr58 (cyan) in the anionic site. (C) Superimposition of the amino acids essential for binding the epalrestat in Tm1743 (cyan, PDB 6KIY), AKR1B1 (yellow, PDB 4JIR), and AKR1B10 (pink, PDB 4JIH). The epalrestat molecules are shown in stick models and colored yellow in Tm1743, dark green in AKR1B1, and hot pink in AKR1B10. The orientation of the epalrestat hydrophobic ring in AKR1B10 is different from that in Tm1743 and AKR1B1. The side chain orientations of Tyr57 and Trp118 in Tm1743 (cyan) are different from those in AKR1B10 (pink). (D) Side chain orientations of the amino acids for immobilizing tolrestat (blue), epalrestat (cyan) in Tm1743, AKR1B1 (yellow), and AKR1B10 (pink). Tyr57 residue replaced the Val in AKR1B1 and AKR1B10. The side chain of Tyr57 is rotated into the anionic site in tolrestat-bound Tm1743 (blue) relative to its position in the epalrestat-bound (cyan) structure. The side chain orientation of Trp118 in tolrestat-bound Tm1743 (blue) is similar to that in epalrestat-bound AKR1B10 (pink).

binding, β5 shifted toward the anionic site, wherein the tyrosyl group of Tyr57 was rotated approximately 115° and inserted into the anionic site to form a hydrogen bond with the carboxyl group of tolrestat. In this bound state, the imidazole ring of His117 and the indole ring of Trp118 were flipped 90° and 180°, respectively (Fig. 5B). The repositioning of the side chains in these conformational changes broadened the anionic site of Tm1743 for trapping the hydrophilic tail of tolrestat, which moved into and occupied the entire anionic site. If $NADP^+$ was manually docked in the tolrestat-bound structure, a 1.8 Å distance between the C15 atom of tolrestat and the C3N atom of the nicotinamide ring would not allow binding of the nicotinamide ring in the anionic site (Fig. 5C).

Other conformational changes were observed in regions corresponding to loop B that are involved in accommodating the pyrophosphate group and adenine ring of $NADP^+$ in both mammalian AKRs and Tm1743 (Figs 2D, 3A and 5A). In $NADP^+$-bound Tm1743, the pyrophosphate group of $NADP^+$ was sandwiched by hydrogen bond interactions between the Arg203 and Ser199 side chains and the main chain nitrogen atoms of Lys244, Leu201, and Pro200 [32]. Whereas in the tolrestat-bound structure, the side chains of Arg202, Arg203, and Lys244 were dramatically shifted or flipped 180°, resulting in movement of the β8-α8 loop toward the $NADP^+$ binding pocket. In particular, the drastic side chain flip of Arg203 uncovered the ribose binding region but closed the cleft that accommodates the adenine ring of $NADP^+$ (Fig. 5B–D

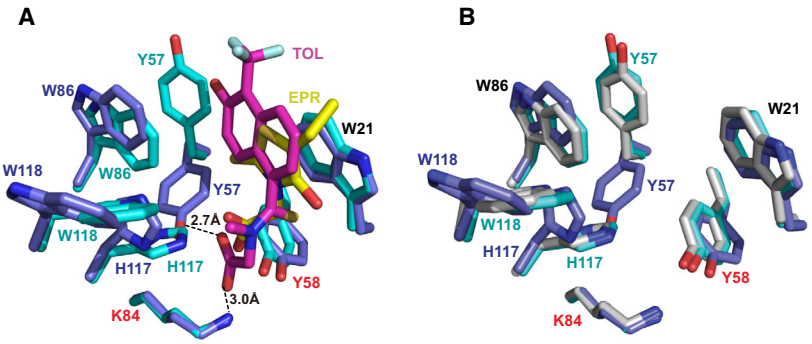

Fig. 4. Comparison of the tolrestat and epalrestat binding in Tm1743. (A) Conformational changes of the residues binding the tolrestat (blue, PDB 6KIK) and epalrestat (cyan, PDB 6KIY) in Tm1743. Upon tolrestat binding, the side chain orientations of Tyr57, Trp118, and His117 are dramatically shifted. The hydrophilic tail of tolrestat (magenta) moves down below the position of epalrestat (yellow) in the anionic site, where it is hydrogen-bonded with Tyr57 and Lys84. (B) Side chain orientations of the essential amino acids in Tm1743 when bound with tolrestat, epalrestat, and $NADP^+$. The side chain orientations of Tyr57, Trp118, and His117 in the tolrestat-bound structure (blue) are different from those in the epalrestat-$NADP^+$- (cyan) and $NADP^+$-bound (gray, PDB 5DAN) Tm1743.

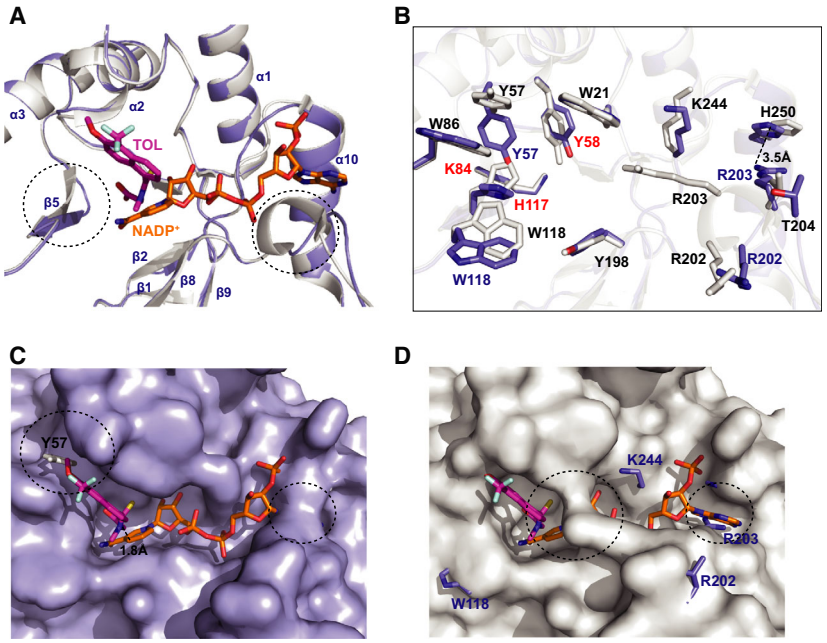

Fig. 5. Comparison of the tolrestat- and $NADP^+$-bound Tm1743. (A) Superimposition of the tolrestat- (blue, PDB 6KIK) and $NADP^+$- (gray, PDB 5DAN) bound Tm1743 pocket (ribbon diagram). The conformational changes to the β5 and β8-α8 regions are indicated with dashed circles. (B) Stick model showing the conformational changes at the anionic site and the cofactor binding pocket of Tm1743 induced by binding with tolrestat (blue) and $NADP^+$ (gray). Flipped side chains of Tyr57, His117, and Trp118 in the anionic site, and Arg202, Arg203, and Thr204 in the adenosine ring binding region are shown in contrast to their positions in the $NADP^+$-bound structure. (C) 3D space-filling models of the tolrestat-bound Tm1743 docked with $NADP^+$ molecule (orange) to show the conformational changes (dashed circles) at the tolrestat and cofactor binding pocket. A 1.8 Å distance was observed between the C15 atom of tolrestat (magenta) and the C3N atom of the $NADP^+$ nicotinamide ring. (D) 3D space-filling models of the $NADP^+$-bound Tm1743 docked with tolrestat (magenta) to show the conformational changes (dashed circles) at the regions necessary for coordinating the nicotinamide and adenosine rings of $NADP^+$ (orange).

and Fig. S6C–D). The amine side chain of Arg203 was thus positioned directly against the imidazole ring of His250, 3.5 Å apart; in this conformation, the Arg203 side chain comes in immediate clash with the adenosine ring of $NADP^+$ (Fig. 5B,D).

Therefore, tolrestat binding resulted in dramatic conformational changes at the anionic site and cofactor binding pockets of Tm1743, which dismantled the anionic site as well as the optimal local environment for accommodating $NADP^+$.

Enzymatic assays confirmed tolrestat as a competitive inhibitor of Tm1743

To verify the competitive binding mode and inhibitory effect of tolrestat, we spectrophotometrically measured the enzymatic reduction of soluble ethyl acetoacetate by Tm1743 and subsequent depletion of NADPH at 340 nm. Addition of tolrestat and epalrestat in the standard reaction system showed a clear inhibitory effect on the overall enzyme activity of Tm1743 (Fig. 6A). As the tolrestat concentration increased, the double-reciprocal plots of the reaction velocity and the substrate concentrations showed an increase in the slopes and an intersection at the *y*-axis (Fig. 6B). Compared to Tm1743 without inhibitor, the apparent Michaelis–Menten constant K_m^{app} increased but the V_m^{app} stay stable as the tolrestat concentrations increased (Fig. S8A and Table 2). These catalytic features indicate that tolrestat acts as a competitive inhibitor of Tm1743 in reducing ethyl acetoacetate. Although the double-reciprocal plots and kinetic constants under different epalrestat concentrations showed a concentration-dependent inhibition, it could not firmly conclude a typical competitive inhibition of epalrestat on the Tm1743 enzyme activity (Fig. S8B and Table 2). By assaying a range of inhibitor concentrations, the apparent inhibitory constant (K_i) of tolrestat and epalrestat was determined to be 87.0 and 55.4 nM, respectively (Fig. 6C and Fig. S8C).

To further elucidate the structural analyses that tolrestat dismantled the $NADP^+$ binding, we measured the enzyme activity of Tm1743 by increasing NADPH concentrations but keep tolrestat at a constant concentration. In the absence of tolrestat, the initial reaction velocities in reducing high concentrations of ethyl acetoacetate were similar at different NADPH concentrations (Fig. S8D). In presence of 25 nM tolrestat, the

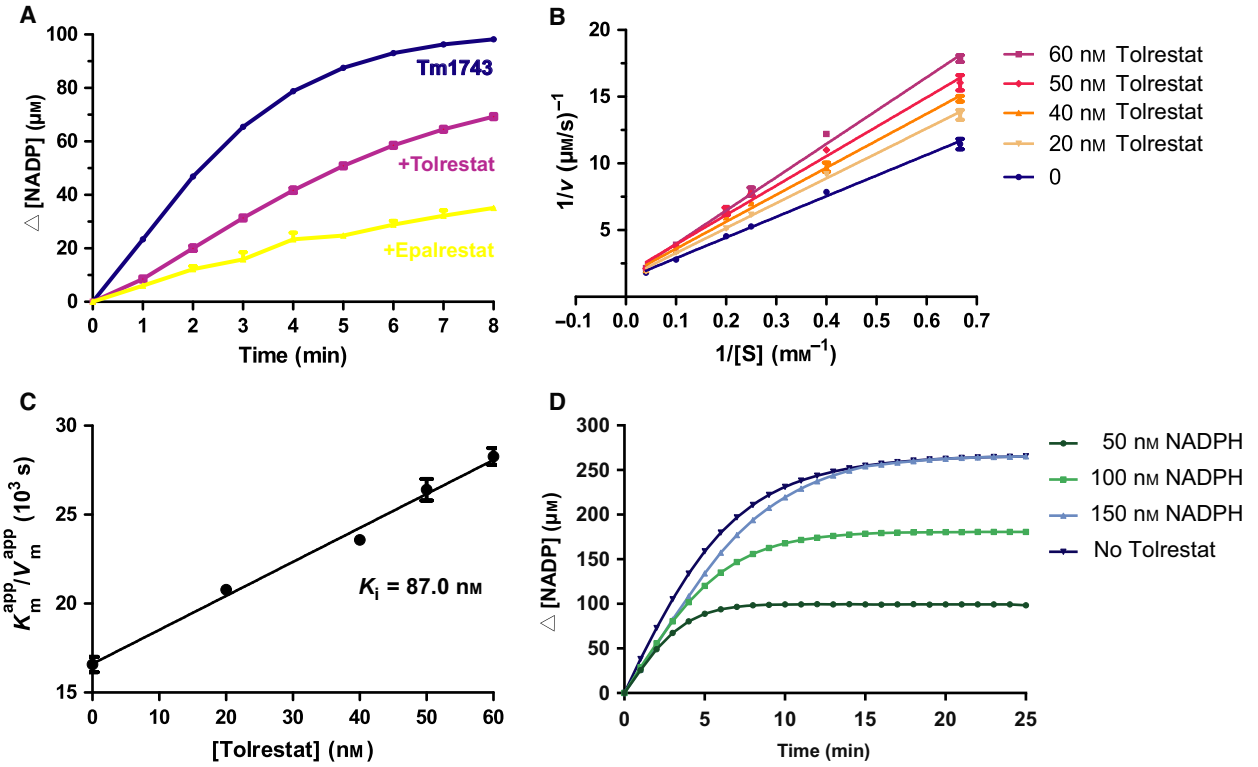

Fig. 6. Enzymatic analyses of the inhibitory effects of tolrestat and epalrestat on Tm1743. (A) Addition of tolrestat and epalrestat individually inhibited the enzymatic activity of Tm1743 in reducing ethyl acetoacetate. Changes in the NADP concentrations were measured over 8 min. (B) Lineweaver–Burk double-reciprocal plot of the reaction velocity and substrate concentrations during exposure to tolrestat concentrations at 0, 20.0, 40.0, 50.0, and 60.0 nM. As the tolrestat concentration increased, the linear regression with increased slopes intersected at one point of the *y*-axis. (C) The apparent kinetic constants K_m^{app}/V_m^{app} were plotted against different inhibitor concentrations to calculate the K_i of tolrestat. The values of K_i are calculated from the *x* intercept. (D) The enzyme activity of Tm1743 was measured by keeping tolrestat concentration at 25 nM and NADPH concentrations at 50.0, 100.0, and 150.0 nM. Changes in the NADP concentrations were measured over 25 min. All the data were obtained from triplicate experiments. Error bars represent SEM as indicated.

initial reaction velocity was decreased but also not affected by the increased NADPH concentrations (Fig. 6D), indicating that NADPH is not able to displace the bound tolrestat molecule from Tm1743 enzyme. These analyses confirmed tolrestat as a competitive inhibitor of Tm1743 and verified our structural observations that tolrestat binding triggered dramatic conformational changes at the cofactor binding pocket of Tm1743 which prevent NADPH binding (Figs 4 and 5).

Binding features characterized from molecular dynamics simulation

To further explore the binding characteristics of tolrestat and epalrestat with Tm1743, MD simulations were performed on the tolrestat-, epalrestat-, and $NADP^+$-bound Tm1743 structures. No drastic fluctuations occurred at the RMSD values during the MD simulations, indicating the conformations of these binding structures are rather stable (Fig. S9A–C). The binding free energies of tolrestat ($\Delta G_{bind} = -8.31$ kcal·mol^{-1}) and epalrestat ($\Delta G_{bind} = -8.83$ kcal·mol^{-1}) agreed well with our experimental results in both the binding strength order (epalrestat > tolrestat) and the values (tolrestat $\Delta G_{bind,exp} = -9.69$ kcal·mol^{-1}, epalrestat $\Delta G_{bind,exp} = -9.96$ kcal·mol^{-1}), validating the reliability of our MD simulations (Table 3). Compared to the inhibitors, $NADP^+$ was more favorable for binding to Tm1743, as implied by the much lower binding free energy (Table 3). However, the binding free energy of $NADP^+$ became less negative upon epalrestat binding, which supports our conclusion from the structural and enzymatic analyses that epalrestat concomitantly binds to Tm1743 with $NADP^+$, and noncompetitively inhibits the enzyme activity of Tm1743 (Fig. 3A,B and Fig. S8B).

The binding free energy was further decomposed to identify the per-residue contribution (Fig. 7). As shown in Fig. 7A, residues that involved in hydrogen bonding interactions with the hydrophilic tail of tolrestat, including Tyr57 and Lys84, displayed major favorable contributions to the binding. The residues Trp21 and Trp86 for stacking to the hydrophobic ring, and Tyr58 and His117 for forming the anionic site also contribute to the tolrestat binding (Fig. 2B,C). However, the catalytic residue Asp53 that did not undergo conformational changes is energetically unfavorable for tolrestat binding, which most likely because of the electrostatic repulsions between the two carboxyl groups (Figs 1B and 7A).

For epalrestat, the major favorable contributions were from the catalytic residue His117 that forms hydrogen bond with the hydrophilic tail, and Trp21 that parallel stacks with the hydrophobic ring (Figs 1D and 3B). Compared to tolrestat, the binding free energy of Tyr57 was decreased, since the hydrogen bond was disrupted in epalrestat binding (Fig. 3B). Notably, $NADP^+$ also displayed favorable contributions to the binding free energy of epalrestat, indicating $NADP^+$ is favorable to the epalrestat binding, which is consistent with their concomitant binding in the crystal structures. On the other side, epalrestat showed minor unfavorable contribution to the $NADP^+$ binding (Fig. 7C,D), which is in accordance with the enzymatic analyses and binding free energy calculations (Fig. 6A and Table 3). Therefore, epalrestat acts as a noncompetitive inhibitor of Tm1743, the concomitant binding with $NADP^+$ is necessary for its inhibition on the enzyme activity.

Discussion

The AKR superfamily of NADP(H)-dependent aldehyde and ketone reductases has been exploited extensively as drug targets in pharmaceutical and clinical trials. AKR inhibitors such as tolrestat and epalrestat have been developed to target AKR1B1 for treatment of complications that accompany type 2 diabetes. However, clinical application of these inhibitors has been limited due to their adverse off-target effects of cross-inhibition on other AKRs, such as AKR1A1, which shares a 65% amino acid sequence identity with AKR1B1 [26,27]. Therefore, in order to successfully develop effective AKR-specific inhibitors, it is

Table 3. Binding free energies between Tm1743 and the inhibitors and cofactor.

	Energy components[a]							
Substrates	ΔE_{ele}	ΔE_{vdW}	ΔG_{GB}	ΔG_{SA}AA	ΔH	$-T\Delta S$	ΔG_{bind}	$\Delta G_{bind,exp}$[b]
Tolrestat[c]	69.72	−23.76	−64.57	−4.07	−22.64	14.33	−8.31	−9.69
Epalrestat[d]	57.84	−16.98	−57.26	−2.59	−18.99	10.16	−8.83	−9.96
$NADP^+$[d]	−335.97	−73.54	307.97	−9.29	−110.83	37.13	−73.70	–
$NADP^+$[e]	−355.50	−82.20	319.85	−9.22	−127.07	38.41	−88.66	–

[a] Energies are in kcal·mol^{-1}.; [b] $\Delta G_{bind,exp} = RT{\cdot}\ln K_{dissociated}$.; [c] Tolrestat was independently bound to Tm1743.; [d] Epalrestat and $NADP^+$ were concomitantly bound to Tm1743.; [e] $NADP^+$ was independently bound to Tm1743.

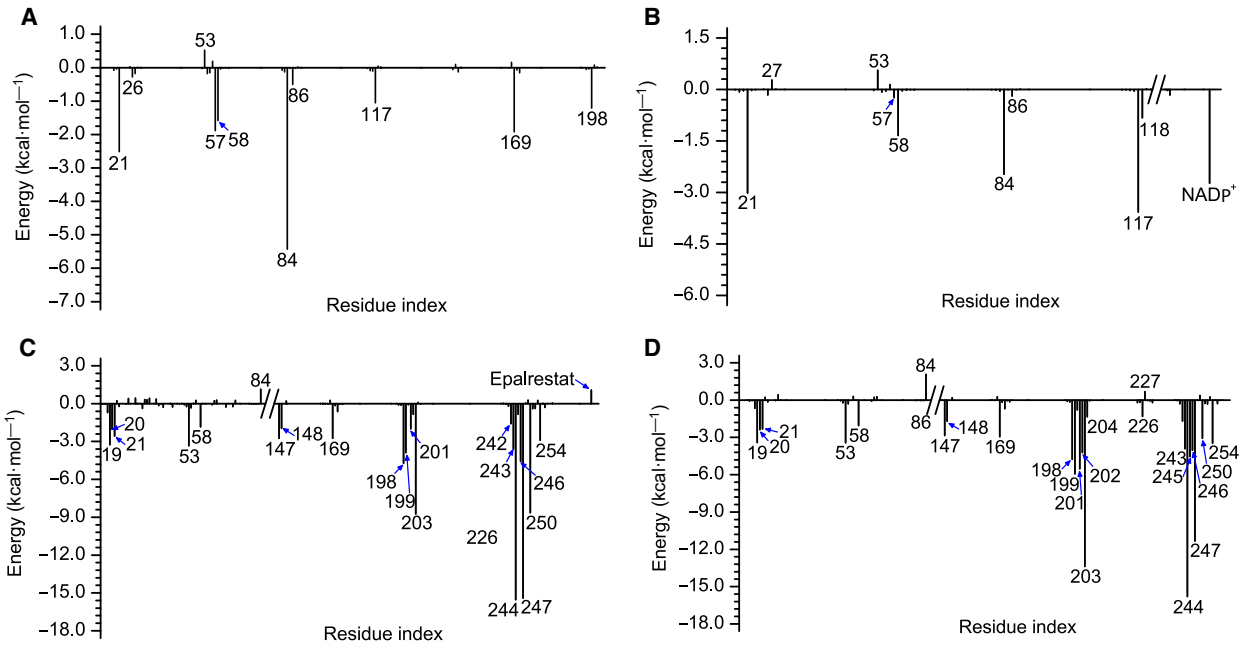

Fig. 7. The per-residue decomposition of the binding free energy of Tm1743 in binding with tolrestat (A), epalrestat (B), epalrestat-bound $NADP^+$ (C), and $NADP^+$ alone (D). The binding free energy values ($kcal \cdot mol^{-1}$) are plotted against the amino acid numbers to show their contributions.

necessary to improve the inhibitor selectivity. We previously investigated the enantioselectivity of Tm1743, a native AKR from *T. maritima*, which was characterized as an efficient biocatalyst in chiral alcohol production. Here, in this work, we determined the crystal structure of Tm1743 in complex, individually, with the two AKR1B1 inhibitors tolrestat and epalrestat. Through structural and enzymatic analyses, we found that tolrestat competitively binds Tm1743 as part of its inhibitory effect. Furthermore, tolrestat binding triggered dramatic conformational changes at the anionic site and the cofactor binding pocket that directly blocked $NADP^+$ binding.

We found three amino acids that were responsible for tolrestat binding in the anionic pocket and that the conformations of these three amino acids differed from their spatial dispositions in other AKRs. Specifically, in mammalian AKRs, three extensions appeared in the loops connecting α4-β5 (loop A), β8-α8 (loop B), and the C terminus of Tm1743, which lifted tolrestat up into a hydrophobic cleft above the anionic site. Therein, the hydrophilic tail of tolrestat is hydrogen-bonded with conserved Trp, His, and Tyr residues (Fig. 2E,F and Fig. S4E,F). Due to lack of these three insertions in Tm1743, the hydrophobic ring of tolrestat is instead sandwiched between the indole rings of strictly conserved Trp86 and Trp21, which typically accommodate the backbone of the tolrestat hydrophilic tail in mammalian AKRs (Fig. 2F and Fig. S4F). In our previous research, Trp21 and Trp86 were shown to play preeminent roles in determining the enantioselectivity of Tm1743 [32]. In the simulated complex structure, the phenyl rings of ethyl 2-oxo-4-phenylbutyrate (EOPB) are located in a hydrophobic pocket formed by the tyrosyl group of Tyr57 and the tryptophanyl groups of Trp21 and Trp86 [32]. Thus, these three amino acids are essential for immobilizing the hydrophobic portion of the substrate.

This study also found that three side chains belonging to Tyr57, Trp118, and His117 are repositioned by tolrestat binding, and this conformational change precludes accommodation of $NADP^+$, thus determining the competitive binding mode for tolrestat. In tolrestat-bound Tm1743, Tyr57 replaces the Val in mammalian AKRs, the tyrosyl group of which is rotated 115° and inserted into the anionic site, where it forms a hydrogen bond with the tolrestat carboxyl group, subsequently instigating a reorientation of the Trp118 and His117 side chains (Fig. 4A,B). These side chain conformational changes of Tyr57, His117, and His118 were consistent with their binding free energy differences between the tolrestat- and epalrestat-bound Tm1743 (Fig. 7A,B). The side chain flip of Trp118 (or its functional counterpart, Trp112) not only broadens the anionic site, but

also abolishes the hydrogen bonding interactions that would otherwise stabilize the hydrophilic tail of tolrestat or epalrestat (Fig. 2C,F, Fig. S7B,D).

Trp118 is also a highly conserved amino acid; its counterparts, Trp112 or Trp111 in mammalian AKRs, were reported to be essential for determining the inhibitor selectivity of AKR1B1 and AR1B10 through a side chain flip that controls the opening and closing of the anionic site [31]. The side chain orientation of Trp118 in tolrestat-bound Tm1743 is similar to that of Trp112 in epalrestat- but not tolrestat-bound AKR1B10 (Fig. 3D), wherein the epalrestat molecule also adopts a different conformation comparable to that in AKR1B1 (Fig. S7A–D). These structural features indicate that the side chain orientation of Tyr57, Trp118, and His117 in the anionic site is a decisive factor in determining the inhibitor-binding mode of AKRs.

Based on the exclusion of $NADP^+$ from its binding site when tolrestat is bound (Figs 2B and 6D), we speculate that occupation of the $NADP^+$ binding pocket is related to the specificity and binding mode of the inhibitor. In other $NADP^+$-bound AKRs, the hydrophilic tails of the inhibitors are located above the anionic site, allowing binding of the $NADP^+$ cofactor. The hydrogen bond formed between the hydrophilic tail and the catalytic residue Lys84 ensures a much stronger tolrestat binding, as implied by its major contribution in the binding free energy of tolrestat (Fig. 7A). Therefore, occupation of the $NADP^+$ binding region is ostensibly the likeliest cause of inhibitory binding by tolrestat in Tm1743, and capacity to bind tolrestat in this region can be exploited for improvement of specificity in the design of mammalian AKR inhibitors. For example, we could plausibly modify inhibitors to switch from the typical noncompetitive binding mode to the competitive mode by elongating the hydrophilic tail, resulting in its insertion into the anionic site and subsequent occupation of the $NADP^+$ binding pocket.

In contrast, the side chain reorientation of Tyr57 in tolrestat-bound Tm1743 provides a new amino acid site for semirational engineering of Tm1743. We previously identified the importance of Trp21, Trp86, and His118 in the enantioselectivity of Tm1743 [32]. Combinations of mutations in these three residues not only changed the conformation of the bound substrate, but also improved the optical purity ratio of the products. Therefore, incorporation of the Tyr57 mutation in the engineering of Tm1743 would benefit the synthesis of optically pure alcohol intermediates for pharmaceutical industry applications. For reactions that require $NADP^+$ as a cofactor, mutation of Tyr57 to a Val or other neutral amino acid in mammalian AKRs may result in a noncompetitive binding mode for different aldehyde and ketone substrates.

The different binding modes and inhibitory effects of tolrestat and epalrestat in the same enzyme, as well as our previous enantioselectivity studies, indicate the flexibility of Tm1743 structure and suggest its potential application as a natural biocatalyst useful for inhibitor design and synthesis of pharmaceuticals. Our work has therefore revealed a previously unrecognized competitive binding mode of tolrestat and illustrated the structural basis for inhibitor and substrate binding in AKRs. These findings will contribute to AKR-targeted drug development and open a wide range of potential pharmaceutical applications for Tm1743 in the future.

Acknowledgements

We thank Professor Zhenmin Chen for providing the vector of Tm1743. We thank staff of the beamlines BL17U1, BL18U1, and BL19U1 at Shanghai Synchrotron Radiation Facility for assistance during data collection. This work was supported by the National Natural Science Foundation of China under grant numbers 31870740, 31570738, and 31400630 to XX. And the Natural Science Foundation of Zhejiang province under grant numbers LY18H250002 to ZW and LY14C050002 to XX.

Author contributions

CZ performed purification, crystallization, data collection and structure determination; ZM and XL performed enzymatic assays; CW determined and refined the structure; ZW performed molecular dynamics simulations; JS performed the gel filtration and AUC analyses; WT, XZ, and DL assisted in all the experiments; and XX designed the experiment and wrote the manuscript. All authors have given approval to the final version of the manuscript.

Data Availability Statement

The structure factors and coordinates of tolrestat-bound and epalrestat-bound Tm1743 have been deposited in the Protein Data Bank under the accession codes 6KIK, 6KIY, and 6KY6.

References

1 Penning TM (2014) The aldo-keto reductases (AKRs): overview. *Chem Biol Int* **234**, 236–246.
2 Mindnich RD and Penning TM (2009) Aldo-keto reductase (AKR) superfamily: genomics and annotation. *Hum Genomics* **3**, 362–370.

3 Hahn-Hagerdal B, Linden T, Senac T and Skoog K (1991) Ethanolic fermentation of pentoses in lignocellulose hydrolysates. *Appl Biochem Biotechnol* **28–29**, 131–144.

4 Anderson S, Marks CB, Lazarus R, Miller J, Stafford K, Seymour J, Light D, Rastetter W and Estell D (1985) Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified *Erwinia herbicola*. *Science* **230**, 144–149.

5 Ikeda H, Nonomiya T, Usami M, Ohta T and Omura S (1999) Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc Natl Acad Sci USA* **96**, 9509–9514.

6 Takikawa H, Ookhtens M, Stolz A and Kaplowitz N (1987) Cyclical oxidation-reduction of the C3 position on bile acids catalyzed by 3 alpha-hydroxysteroid dehydrogenase. II. Studies in the prograde and retrograde single-pass, perfused rat liver and inhibition by indomethacin. *J Clin Invest* **80**, 861–866.

7 Barski OA, Tipparaju SM and Bhatnagar A (2008) The aldo-keto reductase superfamily and its role in drug metabolism and detoxification. *Drug Metab Rev* **40**, 553–624.

8 Chen WD and Zhang Y (2012) Regulation of aldo-keto reductases in human diseases. *Front Pharmacol* **3**, 35.

9 Gabbay KH (1975) Hyperglycemia, polyol metabolism, and complications of diabetes mellitus. *Annu Rev Med* **26**, 521–536.

10 Murata M, Ohta N, Sakurai S, Alam S, Tsai J, Kador PF and Sato S (2001) The role of aldose reductase in sugar cataract formation: aldose reductase plays a key role in lens epithelial cell death (apoptosis). *Chem Biol Interact* **130–132**, 617–625.

11 Liu H, Luo Y, Zhang T, Zhang Y, Wu Q, Yuan L, Chung SS, Oates PJ and Yang JY (2011) Genetic deficiency of aldose reductase counteracts the development of diabetic nephropathy in C57BL/6 mice. *Diabetologia* **54**, 1242–1251.

12 Judzewitsch RG, Jaspan JB, Polonsky KS, Weinberg CR, Halter JB, Halar E, Pfeifer MA, Vukadinovic C, Bernstein L, Schneider M *et al.* (1983) Aldose reductase inhibition improves nerve conduction velocity in diabetic patients. *N Engl J Med* **308**, 119–125.

13 Penning TM, Burczynski ME, Jez JM, Hung CF, Lin HK, Ma H, Moore M, Palackal N and Ratnam K (2000) Human 3alpha-hydroxysteroid dehydrogenase isoforms (AKR1C1-AKR1C4) of the aldo-keto reductase superfamily: functional plasticity and tissue distribution reveals roles in the inactivation and formation of male and female sex hormones. *Biochem J* **351**, 67–77.

14 Rizner TL, Smuc T, Rupreht R, Sinkovec J and Penning TM (2006) AKR1C1 and AKR1C3 may determine progesterone and estrogen ratios in endometrial cancer. *Mol Cell Endocrinol* **248**, 126–135.

15 Byrns MC, Duan L, Lee SH, Blair IA and Penning TM (2010) Aldo-keto reductase 1C3 expression in MCF-7 cells reveals roles in steroid hormone and prostaglandin metabolism that may explain its over-expression in breast cancer. *J Steroid Biochem Mol Biol* **118**, 177–187.

16 Gonzales E, Cresteil D, Baussan C, Dabadie A, Gerhardt MF and Jacquemin E (2004) SRD5B1 (AKR1D1) gene analysis in delta(4)-3-oxosteroid 5beta-reductase deficiency: evidence for primary genetic defect. *J Hepatol* **40**, 716–718.

17 Drury JE, Mindnich R and Penning TM (2010) Characterization of disease-related 5beta-reductase (AKR1D1) mutations reveals their potential to cause bile acid deficiency. *J Biol Chem* **285**, 24529–24537.

18 Gallego O, Belyaeva OV, Porte S, Ruiz FX, Stetsenko AV, Shabrova EV, Kostereva NV, Farres J, Pares X and Kedishvili NY (2006) Comparative functional analysis of human medium-chain dehydrogenases, short-chain dehydrogenases/reductases and aldo-keto reductases with retinoids. *Biochem J* **399**, 101–109.

19 Gallego O, Ruiz FX, Ardevol A, Dominguez M, Alvarez R, de Lera AR, Rovira C, Farres J, Fita I and Pares X (2007) Structural basis for the high all-trans-retinaldehyde reductase activity of the tumor marker AKR1B10. *Proc Natl Acad Sci USA* **104**, 20764–20769.

20 Ma J, Luo DX, Huang C, Shen Y, Bu Y, Markwell S, Gao J, Liu J, Zu X, Cao Z *et al.* (2012) AKR1B10 overexpression in breast cancer: association with tumor size, lymph node metastasis and patient survival and its potential as a novel serum marker. *Int J Cancer* **131**, E862–E871.

21 Sestanj K, Bellini F, Fung S, Abraham N, Treasurywala A, Humber L, Simard-Duquesne N and 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.

22 Hotta N, Kawamori R, Fukuda M, Shigeta Y and Aldose Reductase Inhibitor-Diabetes Complications Trial Study Group (2012) Long-term clinical effects of epalrestat, an aldose reductase inhibitor, on progression of diabetic neuropathy and other microvascular complications: multivariate epidemiological analysis based on patient background factors and severity of diabetic neuropathy. *Diabet Med* **29**, 1529–1533.

23 Larson ER, Lipinski CA and Sarges R (1988) Medicinal chemistry of aldose reductase inhibitors. *Med Res Rev* **8**, 159–186.

24 Humber LG (1987) The medicinal chemistry of aldose reductase inhibitors. *Prog Med Chem* **24**, 299–343.

25 Scuric Z, Stain SC, Anderson WF and Hwang JJ (1998) New member of aldose reductase family proteins overexpressed in human hepatocellular carcinoma. *Hepatology* **27**, 943–950.

26 Pastel E, Pointud JC, Volat F, Martinez A and Lefrancois-Martinez AM (2012) Aldo-keto reductases 1B in endocrinology and metabolism. *Front Pharmacol* **3**, 148.
27 Petrash JM (2004) All in the family: aldose reductase and closely related aldo-keto reductases. *Cell Mol Life Sci* **61**, 737–749.
28 Urzhumtsev A, Tete-Favier F, Mitschler A, Barbanton J, Barth P, Urzhumtseva L, Biellmann JF, Podjarny A and Moras D (1997) A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. *Structure* **5**, 601–612.
29 Steuber H, Zentgraf M, Gerlach C, Sotriffer CA, Heine A and Klebe G (2006) Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions. *J Mol Biol* **363**, 174–187.
30 Steuber H, Heine A, Podjarny A and Klebe G (2008) Merging the binding sites of aldose and aldehyde reductase for detection of inhibitor selectivity-determining features. *J Mol Biol* **379**, 991–1016.
31 Zhang L, Zhang H, Zhao Y, Li Z, Chen S, Zhai J, Chen Y, Xie W, Wang Z, Li Q *et al.* (2013) Inhibitor selectivity between aldo-keto reductase superfamily members AKR1B10 and AKR1B1: role of Trp112 (Trp111). *FEBS Lett* **587**, 3681–3686.
32 Wang Z, Zhou S, Zhang S, Zhang S, Zhu F, Jin X, Chen Z and Xu X (2017) Semi-rational engineering of a thermostable aldo-keto reductase from *Thermotoga maritima* for synthesis of enantiopure ethyl-2-hydroxy-4-phenylbutyrate (EHPB). *Sci Rep* **7**, 4007.
33 Ma Y-H, Lv D-Q, Zhou S, Lai D-Y and Chen Z-M (2013) Characterization of an aldo–keto reductase from *Thermotoga maritima* with high thermostability and a broad substrate spectrum. *Biotechnol Lett* **35**, 757–762.
34 Minor W, Cymborowski M, Otwinowski Z and Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62**, 859–866.
35 Vaguine AA, Richelle J and Wodak SJ (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191–205.
36 Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491–497.
37 Collaborative Computational Project, N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.
38 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC and Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
39 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
40 Murshudov GN, Vagin AA and Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
41 Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B and Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem.* **26**, 1668–1688.
42 Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE and Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* **11**, 3696–3713.
43 Cummins PL, Ramnarayan K, Singh UC and Gready JE (1991) Molecular dynamics/free energy perturbation study on the relative affinities of the binding of reduced and oxidized NADP to dihydrofolate reductase. *J Am Chem Soc* **113**, 8247–8256.
44 Bayly CI, Cieplak P, Cornell W and Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* **97**, 10269–10280.
45 Wang Z, Chen Z, Li J, Huang J, Zheng C and Liu JP (2019) Combined 3D-QSAR, molecular docking and molecular dynamics study on the benzimidazole inhibitors targeting HCV NS5B polymerase. *J Biomol Struct Dyn* **27**, 1–12.
46 Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W *et al.* (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* **33**, 889–897.
47 Liu N, Zhou W, Guo Y, Wang J, Fu W, Sun H, Li D, Duan M and Hou T (2018) Molecular dynamics simulations revealed the regulation of ligands to the interactions between androgen receptor and its coactivator. *J Chem Inf Model* **58**, 1652–1661.
48 Weiser J, Shenkin PS and Still WC (1999) Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities. *Biopolymers* **50**, 373–380.
49 Yang Y, Shen Y, Liu H and Yao X (2011) Molecular dynamics simulation and free energy calculation studies of the binding mechanism of allosteric inhibitors with p38alpha MAP kinase. *J Chem Inf Model* **51**, 3235–3246.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. The electron density of tolrestat (A), epalrestat and cofactor $NADP^+$ (B) resolved in Tm1743.
Fig. S2. Overall structure of epalrestat-bound Tm1743 with $P3_221$ space group (PDB 6KY6).
Fig. S3. Gel filtration and analytical ultracentrifugation (AUC) of Tm1743.
Fig. S4. Comparison of the tolrestat binding pocket in Tm1743 with mammalian AKR1B and AKR1B10.
Fig. S5. Structure-based sequence alignment of Tm1743 with mammalian AKR1B1 and AKR1B10.
Fig. S6. Comparison of the tolrestat and cofactor binding pocket in Tm1743 with mammalian AKRs.
Fig. S7. The epalrestat binding pocket in mammalian AKRs.
Fig. S8. Enzymatic characterization of Tm1743 in reducing ethyl acetoacetate.
Fig. S9. Root-mean-square deviations (RMSDs) of Tm1743 and the bound inhibitors and $NADP^+$ during MD simulations.