Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structures of AzrA and of AzrC complexed with substrate or inhibitor: insight into substrate specificity and catalytic mechanism

Jian Yu,[a] Daiki Ogata,[b] ZuoQi Gai,[a] Seiichi Taguchi,[b] Isao Tanaka,[a] Toshihiko Ooi[b]* and Min Yao[a,c]*

[a]Faculty of Advanced Life Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan, [b]Division of Biotechnology and Macromolecular Chemistry, Graduate School of Engineering, Hokkaido University, Kita 13, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan, and [c]Basic Medical College of Zhengzhou University, Department of Pharmacology, No. 100 Science Avenue, Zhengzhou, Henan 450001, People's Republic of China

Correspondence e-mail: ooi@eng.hokudai.ac.jp, yao@castor.sci.hokudai.ac.jp

Received 9 August 2013
Accepted 11 November 2013

PDB references: AzrA, 3w77; AzrC, complex with Cibacron Blue, 3w78; complex with Orange I, 3w79; complex with Acid Red 88, 3w7a

Azo dyes are major synthetic dyestuffs with one or more azo bonds and are widely used for various industrial purposes. The biodegradation of residual azo dyes *via* azoreductase-catalyzed cleavage is very efficient as the initial step of wastewater treatment. The structures of the complexes of azoreductases with various substrates are therefore indispensable to understand their substrate specificity and catalytic mechanism. In this study, the crystal structures of AzrA and of AzrC complexed with Cibacron Blue (CB) and the azo dyes Acid Red 88 (AR88) and Orange I (OI) were determined. As an inhibitor/analogue of NAD(P)H, CB was located on top of flavin mononucleotide (FMN), suggesting a similar binding manner as NAD(P)H for direct hydride transfer to FMN. The structures of the AzrC–AR88 and AzrC–OI complexes showed two manners of binding for substrates possessing a hydroxy group at the *ortho* or the *para* position of the azo bond, respectively, while AR88 and OI were estimated to have a similar binding affinity to AzrC from ITC experiments. Although the two substrates were bound in different orientations, the hydroxy groups were located in similar positions, resulting in an arrangement of electrophilic C atoms binding with a proton/electron-donor distance of ∼3.5 Å to N5 of FMN. Catalytic mechanisms for different substrates are proposed based on the crystal structures and on site-directed mutagenesis analysis.

1. Introduction

Azo dyes are major synthetic dyestuffs with one or more azo bonds and are widely used in large quantities for various industrial purposes in the printing, textile, food and cosmetics industries because of their versatility, ease of synthesis, chemical stability and range of uses (Griffiths, 1984). Industrial effluents are often contaminated with residual dyes, which not only affect water quality but are also threats to public health. In some cases, these compounds have carcinogenic and mutagenic potential, and are therefore of significant concern as environmental pollutants (Chung *et al.*, 1981; Chung, 1983). Wastewater-treatment systems for azo dye-containing effluent based on physicochemical treatments are efficient but have high running costs and consume chemicals and energy (Kariminiaae-Hamedaani *et al.*, 2007). Hence, the biological treatment of azo dyes using bacteria has been widely studied. Generally, it is assumed that the initial step in the biodegradation of azo dyes is conversion to the amine *via* azoreductase-catalyzed cleavage. Therefore, azoreductases have attracted attention as key enzymes for bioremediation (Stolz, 2001). In addition to bioremediation, azoreductases from intestinal bacteria have also attracted attention for site-specific activation of the azo pro-drugs balsalazide and

sulfasalazine in the treatment of inflammatory bowel disease (Dissanayake & Truelove, 1973; Hanauer, 1996). It has been suggested that AzoR from *Escherichia coli* is involved in protection against thiol-specific stress by electrophilic quinones (Liu *et al.*, 2009). However, azo dyes are artificial chemical compounds, and the physiological function of bacterial azoreductases is poorly understood.

These enzymes are classified into at least two different types in bacteria: flavin mononucleotide-containing (FMN) azoreductases and flavin-free azoreductases (Chen, 2006). FMN azoreductases catalyze the reductive cleavage of the azo bond (—N═N—) in azo compounds *via* a ping-pong or ping-pong bi-bi mechanism requiring one or two moles of NAD(P)H to reduce one mole of Methyl Red (Nakanishi *et al.*, 2001), which is a model azo compound, to 2-aminobenzoic acid and *N*,*N*′-dimethyl-*p*-phenylenediamine. The FMN azoreductases can be further subdivided into two types, NADH-dependent or NADPH-dependent azoreductases, based on the selectivity of the coenzymes and their amino-acid sequence similarities (Chen *et al.*, 2005; Liu *et al.*, 2007). Azoreductases belonging to all of these groups have been found in a wide variety of bacteria. For example, FMN azoreductases from *E. coli* (Nakanishi *et al.*, 2001), *Enterococcus faecalis* (Chen *et al.*, 2004), *Bacillus subtilis* (Nishiya & Yamamoto, 2007), *Bacillus* sp. B29 (Ooi *et al.*, 2007, 2009) and *Geobacillus stearothermophilus* (Matsumoto *et al.*, 2010) are NADH-dependent enzymes, while azoreductases from *Bacillus* sp. OY1-2 (Suzuki *et al.*, 2001), *Rhodobacter sphaeroides* (Bin *et al.*, 2004), *Staphylococcus aureus* (Chen *et al.*, 2005), *B. subtilis* (Sugiura *et al.*, 2006) and *G. stearothermophilus* (Sugiura *et al.*, 2006) are NADPH-dependent enzymes.

Previously, we identified and characterized novel NADH-dependent FMN azoreductases from the bacterium *Bacillus* sp. B29, termed AzrA, AzrB and AzrC (Ooi *et al.*, 2007, 2009), which are capable of degrading several azo dyes. These azoreductases share sequence similarities of 62% between AzrA and AzrB, 41% between AzrA and AzrC, and 37% between AzrB and AzrC. Biochemical analyses indicated that all three of the azoreductases from *Bacillus* sp. B29 are homodimeric proteins with a molecular weight of about 48 kDa that contain two moles of noncovalently bound FMN per mole of dimeric protein and display similar catalytic properties following the ping-pong mechanism. They catalyze the reductive cleavage of the azo linkage, with NADH as an electron donor. Compared with AzrB and AzrC, AzrA showed a wide range of substrate specificity towards azo dyes, including sulfone-modified azo dyes such as Orange I [*p*-(4-hydroxy-1-naphthylazo)benzene sulfonic acid] and Orange II [4-(2-hydroxy-1-naphthylazo)-benzene sulfonic acid]. The reasons for the differences in substrate specificity towards sulfone-modified azo dyes are still unclear.

To understand the substrate recognition and catalytic mechanism of azoreductases, detailed knowledge of the structures of azoreductases in complex with substrates are indispensable. The crystal structures of the azoreductases AzoR from *E. coli* (Ito *et al.*, 2005, 2006), AzoA from *E. faecalis* (Liu *et al.*, 2007) and YhdA from *B. subtilis* (Binter *et al.*, 2009), which are homologues of the three isozymes AzrA, AzrB and AzrC, respectively, have recently been reported. In addition, the structure of the azoreductase paAzoR1 from *Pseudomonas aeruginosa* has been reported in complex with a simple azo dye, Methyl Red, and substrate specificity has been discussed based on homology-modelled structures of paAzoR2 and paAzoR3 (Wang *et al.*, 2007; Ryan *et al.*, 2010). However, we still lack the structural information on azoreductases in complex with various substrates that is necessary to gain a full understanding of their substrate specificity and catalytic mechanism. In this study, we provide structural evidence for the substrate specificity and catalytic mechanism of Azrs by solving the crystal structures of AzrA and of AzrC complexed with the inhibitor Cibacron Blue (CB) of NAD(P)H and the sulfone-modified azo dyes Acid Red 88 (AR88) and Orange I (OI). Based on the architectures of the active sites in all of the present structures and site-directed mutagenesis, we will discuss the substrate specificity and catalytic mechanism for the reduction of azo dyes.

2. Materials and methods

2.1. Protein expression and purification

All of the chemical components, including the azo dyes, were of analytical grade or of the highest quality available and were purchased from Wako Pure Chemical Industries (Osaka, Japan), Tokyo Chemical Industry (Tokyo, Japan) or Sigma–Aldrich (St Louis, Missouri, USA). DNA primers for site-directed mutagenesis were synthesized by Sigma–Aldrich. The amino-acid sequences of the type 1 family azoreductases were aligned using *ClustalW* (Thompson *et al.*, 1994) at the DNA Data Bank of Japan.

The recombinant azoreductases from *Bacillus* sp. B29, selenomethionine-labelled AzrA (SeMet-AzrA) and AzrC, were expressed and purified as described previously (Ogata *et al.*, 2010). Site-directed mutagenesis of AzrA (Y148A) and AzrC (Y151A) was performed by the inverse PCR method with KOD -Plus- DNA polymerase (Toyobo, Japan). The following mutagenic oligonucleotide primers were used: 5′-CAGCTGAGGTAGAAATGGCTG-3′ and 5′-CTGGACCT TCTGA**TGC**TACGCCG-3′ as the primer pair for pET3a-*azrA* and 5′-CAGGTGGCGTA**GCA**TCTGAAGGAGCATA-3′ and 5′-TTGCTTGAATGTGAAGTGCTTTTTTACC-3′ as the primer pair for pET3a-*azrC*. Both of the Tyr-Ala mutants were expressed in *E. coli* BL21 (DE3) cells using LB medium with 1 m*M* IPTG induction at 37°C as described previously (Ooi *et al.*, 2007, 2009). Purification of mutants was performed using DEAE-Cellulose, Toyopearl Butyl-650 and Sephacryl S-200 HR based on the method reported previously (Ooi *et al.*, 2007, 2009).

2.2. Crystallization and data collection

Crystallization and data collection of SeMet-AzrA were performed as described previously (Ogata *et al.*, 2010). Two substrates, sulfonated AR88 and OI, were selected as azo compounds for crystallization with AzrC as they have high

Table 1
Summary of crystallographic data.

Values in parentheses are for the highest resolution shell.

	AzrA	AzrC–CB	AzrC–AR88	AzrC–OI
Wavelength (Å)	0.97918	1.00000	1.00000	1.00000
Space group	$P2_12_12_1$	$P3_221$	$C2$	$C2$
Unit-cell parameters (Å, °)	a = 57.3, b = 69.6, c = 105.2	a = b = 146.4, c = 116.2	a = 191.2, b = 56.7, c = 105.1, β = 115.706	a = 190.9, b = 56.6, c = 104.8, β = 115.726
Resolution range (Å)	50.0–1.66 (1.72–1.66)	47.911–2.62 (2.78–2.62)	50.0–2.1 (2.18–2.10)	50.0–2.4 (2.49–2.40)
No. of unique reflections	50324	43144	59046	39733
Average multiplicity	6.8 (6.1)	8.2 (8.0)	5.1 (4.0)	5.5 (5.1)
Completeness (%)	99.8 (99.0)	99.3 (97.4)	98.9 (94.8)	99.5 (98.6)
$\langle I/\sigma(I)\rangle$	22.22 (3.23)	23.8 (3.34)	11.89 (1.97)	9.79 (2.03)
$R_{\rm merge}$† (%)	7.8 (47.6)	7.8 (63.6)	17.1 (49.7)	12.8 (49.5)
Wilson B factor (Å^2)	18.6	67.0	20.8	38.1

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of symmetry-equivalent reflections hkl.

Table 2
Structure-refinement statistics.

	AzrA	AzrC–CB	AzrC–AR88	AzrC–OI
Resolution range (Å)	44.24–1.66	50.00–2.62	17.21–2.10	36.72–2.40
No. of reflections used	46713	39983	54492	37655
Completeness (%)	99.70	99.25	97.19	98.92
$R/R_{\rm free}$† (%)	18.0/21.1	25.3/28.5	19.9/23.9	22.1/26.4
No. of molecules in asymmetric unit	2	4	4	4
No. of non-H atoms				
Protein	3196	6440	6440	6440
Water	307	21	232	39
Others	62	328	239	216
Averaged B factors (Å^2)				
Protein	14.9	29.0	19.9	26.0
Water	24.7	40.5	29.6	25.3
Others	13.0	66.2	19.2	48.0
R.m.s. deviations				
Bond lengths (Å)	0.012	0.005	0.013	0.0012
Bond angles (°)	1.479	1.027	1.579	1.516
Ramachandran plot‡, residues in (%)				
Favoured region	95.8	96.63	97.0	96.63
Allowed region	4.2	3.3	3.0	3.37
Outlier region	0	0.1	0	0.0

† $R = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are observed and calculated structure-factor amplitudes, respectively. $R_{\rm free}$ was calculated as for R but using a subset (10%) of reflections that were not used for refinement. ‡ Calculated using *PROCHECK*.

solubility in water and different molecular sizes. The crystallization of AzrC complexed with the substrates AR88 (AzrC–AR88) or OI (AzrC–OI) was carried out by co-crystallization using the hanging-drop vapour-diffusion method. Crystals of both AzrC–AR88 (red crystals) and AzrC–OI (orange crystals) were grown under the conditions described previously (Ogata *et al.*, 2010). Blue crystals of AzrC complexed with Cibacron Blue (AzrC-CB) were obtained within several months under the following conditions: a protein concentration of 10 mg ml^{-1} with 0.1 *M* cacodylate buffer pH 6.4, 0.24–0.3 *M* calcium acetate, 30–34%(*w*/*v*) PEG 600 and 2 m*M* Cibacron Blue (Tokyo Chemical Industry).

Data sets for native AzrA and AzrC–AR88 were collected on beamline 41XU at SPring-8 (Proposal Nos. 2007B1095 and 2009A1116), while data collection for AzrC–CB and AzrC–OI was carried out on beamline NW12 at the Photon Factory (PF). All of the crystals in this work were flash-cooled directly at 100 K for data collection without any additional cryoprotectant. The data sets for native AzrA, AzrC–AR88 and AzrC–OI were indexed, integrated, scaled and merged using the *HKL*-2000 package (Otwinowski & Minor, 1997), while the data for AzrC–CB were processed with *XDS* (Kabsch, 2010). The crystallographic parameters and data-collection statistics are shown in Table 1.

2.3. Structure determination and refinement

The structure of SeMet-AzrA was solved by the Se-SAD method. All ten Se-atom sites were found using *SHELXD* (Sheldrick, 2008). The phases were calculated and improved, and an initial structure (78%) was then constructed automatically by *SOLVE*/*RESOLVE* (Terwilliger, 2003) at 2.2 Å resolution. Additional model building, fitting and refinement were carried out automatically using *LAFIRE* (Yao *et al.*, 2006) running the refinement program *CNS* (Brünger *et al.*, 1998; Brunger, 2007), followed by several cycles of manual model checking and fitting with the graphics program *Coot* (Emsley *et al.*, 2010). FMN was then automatically built and fitted by *LAFIRE*. Final refinement of AzrA was performed using native data to 1.66 Å resolution with *REFMAC*5 (Murshudov *et al.*, 2011); TLS (Painter & Merritt, 2006) refinement for each molecule in the asymmetric unit was performed in the last several cycles.

While the structures of the AzrC–AR88 and AzrC–OI complexes were solved by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using the refined structure of AzrA as a search model, the AzrC–CB structure was solved using the refined AzrC–AR88 structure excluding AR88 as a search model. The CB, AR88 and OI molecules in the complexes were built based on their respective $2F_{\rm o} - F_{\rm c}$ and $F_{\rm o} - F_{\rm c}$ maps. Finally, TLS refinement with non-crystallographic symmetry restraints was performed using *REFMAC*5 for all structures. *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT IF* (Vriend, 1990) were used to assess the quality of the final models. The refinement statistics are shown in Table 2.

2.4. Enzyme assay

The protein concentration was measured by the Bradford method using a protein assay reagent (Bio-Rad, Hercules, California) with bovine serum albumin (BSA) as a standard (Bradford, 1976). The enzymatic activity assay was carried out using the method described previously (Ooi *et al.*, 2007, 2009). Standard reaction mixtures consisting of 25 m*M* Tris–HCl buffer pH 7.4, 25 µ*M* Methyl Red (MR), 100 µ*M* NADH and a suitable amount of enzyme in a volume of 3 ml were incubated at 30°C. Reaction mixture without MR was pre-incubated for 3 min and the degradation of MR was followed by monitoring the initial rate of the decrease in absorbance at 430 nm using a spectrophotometer (UV-2500PC; Shimadzu, Kyoto, Japan). One unit of enzyme activity was defined as the amount of enzyme required to decolorize 1 µmol of dye per minute under the assay conditions. The kinetic analysis was performed according to the published methods (Ooi *et al.*, 2007, 2009), with slight modifications. The initial velocities of NADH oxidation and MR reduction were measured at constant concentrations of MR (40, 55, 80 or 120 µ*M*) or NADH (20, 25, 35 or 50 µ*M*). The kinetic constants were calculated using a Lineweaver–Burk double-reciprocal plot of initial velocity *versus* each concentration of NADH and MR (Lineweaver & Burk, 1934). The NADH concentration was measured by the absorbance at 340 nm using a molecular extinction coefficient of 5060 M^{-1} cm^{-1}. All values were determined at least in triplicate.

2.5. Binding-affinity analysis

Isothermal titration calorimetry (ITC) was applied to measure the binding affinity of AzrC for AR88 and OI. The experiments were performed on a Nano ITC (TA Instruments, USA). The samples of AR88 and OI were dissolved in 25 m*M* Tris–HCl pH 7.5, the same buffer as was used for purified AzrC with FMN. The concentration of AzrC was measured at 450 nm by the absorption of FMN using an absorption coefficient of 12 500 M^{-1} cm^{-1}, and the concentration of AR88 and OI was calculated by measuring the absorption at 500 and 477 nm, respectively (Blümel *et al.*, 2002; Pedraza *et al.*, 2006). 50 µl of ligand (AR88 at 216 µ*M* or OI at 220 µ*M*) in the syringe was injected as 25 aliquots of 2 µl each into AzrC (37 µ*M*, 170 µl) in the cell. The experiments were carried out at 25°C with an injection interval of 180 s. The same concentrations of ligands were injected into the buffer as a background.

3. Results

3.1. Overall structures of AzrA and AzrC complexes

The structure of AzrA was determined at 2.2 Å resolution by the Se-SAD method and was extended to 1.66 Å resolution using the native data, while the structures of the AzrC complexes were solved by molecular replacement. All of the residues were built except for the first Met residue at the N-terminus of each structure. Similar to the members of the azoreductase family, the structures of both AzrA and AzrC were formed of a dimer, although the crystals of AzrA and the AzrC complexes belonged to different space groups ($P2_12_12_1$, $P3_221$ and *C*2) with different numbers of molecules in the asymmetric unit (Table 2).

Each monomer forms a flavodoxin-like α/β-fold structure and binds a molecule of FMN at the active site between the monomers of a dimer. The α/β structure consists of seven α-helices and one α_{10}-helix around a β-sheet formed by five parallel β-strands and a β–turn–β structure, with topology $\beta1$–$\alpha1$–$\beta2$–$\alpha2$–α_{10}–$\alpha3$–$\beta3$–$\alpha4$–$\beta4$–t–$\beta5$–$\beta6$–$\alpha5$–$\alpha6$–$\beta7$–$\alpha7$ (Fig. 1*a*). Although the sequence identity between AzrA and AzrC is only 41%, the folds of the overall structures of AzrA and AzrC (in the AzrC–AR88 complex) are very similar, with an r.m.s.d. (root-mean-square deviation) of 1.39 Å (171 C^{α} atoms) apart for five loops 11–18, 60–70, 124–129, 145–153 and 185–189 (r.m.s.d of 3.44 Å for 39 C^{α} atoms) (Figs. 1*b* and 1*c*). Interestingly, these five loops, which are around the active site, have different conformations even though the identity (66.7%) in these loops is higher than that of the total residues.

Interestingly, the crystal packing of the AzrC–CB complex showed that two of the four monomers in the asymmetric unit form a dimer and the other two monomers form two dimers with their crystal symmetry partners, respectively. Among the four molecules in the asymmetric unit of AzrC–CB, molecule *A* has high r.m.s.d values (1.084, 1.304 and 1.086 Å) to the other three molecules. Structure comparison shows that the main conformational change occurrs at the Asn12–Glu15 loop, and the longest distance between C^{α} atoms is 10.2 Å in Ala14. In the crystal structure, molecule *A* contacts another crystal symmetric molecule, inducing conformational change of the Asn12–Glu15 loop.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://wwpdb.org; entries 3w77, 3w78, 3w79 and 3w7a).

3.2. Active site of AzrA and AzrC

The sequence and structural differences between AzrA and AzrC made a difference at the active-site pocket. The active-site pocket of AzrA was larger than that of AzrC (Fig. 2). The average volume of the active-site pockets of the two AzrA molecules was 804 [(824 + 784)/2] $Å^3$ as estimated using *POCASA* (Yu *et al.*, 2010) and was 1.7 times that of AzrC [(498 + 434)/2 = 466 $Å^3$]. Three aligned residue pairs, Asn11/Pro11, Val52/Ile52 and Leu102/Phe105, between the AzrA/AzrC enzymes affected the geometric characteristics (depth and size) of the active-site pocket. The different residue types Asn11/Pro11 and Val52/Ile52 make the active-site pocket of AzrA deeper than that of AzrC (Fig. 2*c*). The average depth was 3.8 Å for the AzrA active-site pocket and 3.3 Å for the AzrC active-site pocket as calculated using *POCASA*. On the other hand, Phe105 limits the size of the active-site pocket in AzrC compared with Leu102 in AzrA. The Phe105 residue is conserved in AzrB (Phe102) and in AzoR from *E. coli* (Phe98), and also limits the pocket size in paAzoR1 (Ryan *et al.*, 2010). The activity of AzrB and AzrC towards structurally large substrates is much lower than that of AzrA (unpublished

data), which suggests that Phe at this position makes entering the active-site pocket difficult for large substrates. In addition, the larger active-site pocket may also allow AzrA to specify more types of substrates than AzrC. Such an activity difference between azoreductase enzymes towards different sized substrates is also found in the homologous paAzoR1 enzymes (Ryan *et al.*, 2010).

The conformation of the FMN-binding site was also slightly different between AzrA and AzrC, although the residues around the FMN-binding site are highly conserved. FMN was positioned to form 15 hydrogen bonds to 13 residues in AzrA, while it bound to ten residues with 12 hydrogen bonds in AzrC. Among the hydrogen bonds to FMN in AzrA, 11 hydrogen bonds (70% of the total hydrogen bonds to FMN) were formed by the ribityl side chain and the phosphoric acid basal part of FMN, whereas there were only six hydrogen bonds (50% of the total hydrogen bonds to FMN) in AzrC. Such binding characteristics can be considered to influence the FMN-binding affinity between the two enzymes, such that the

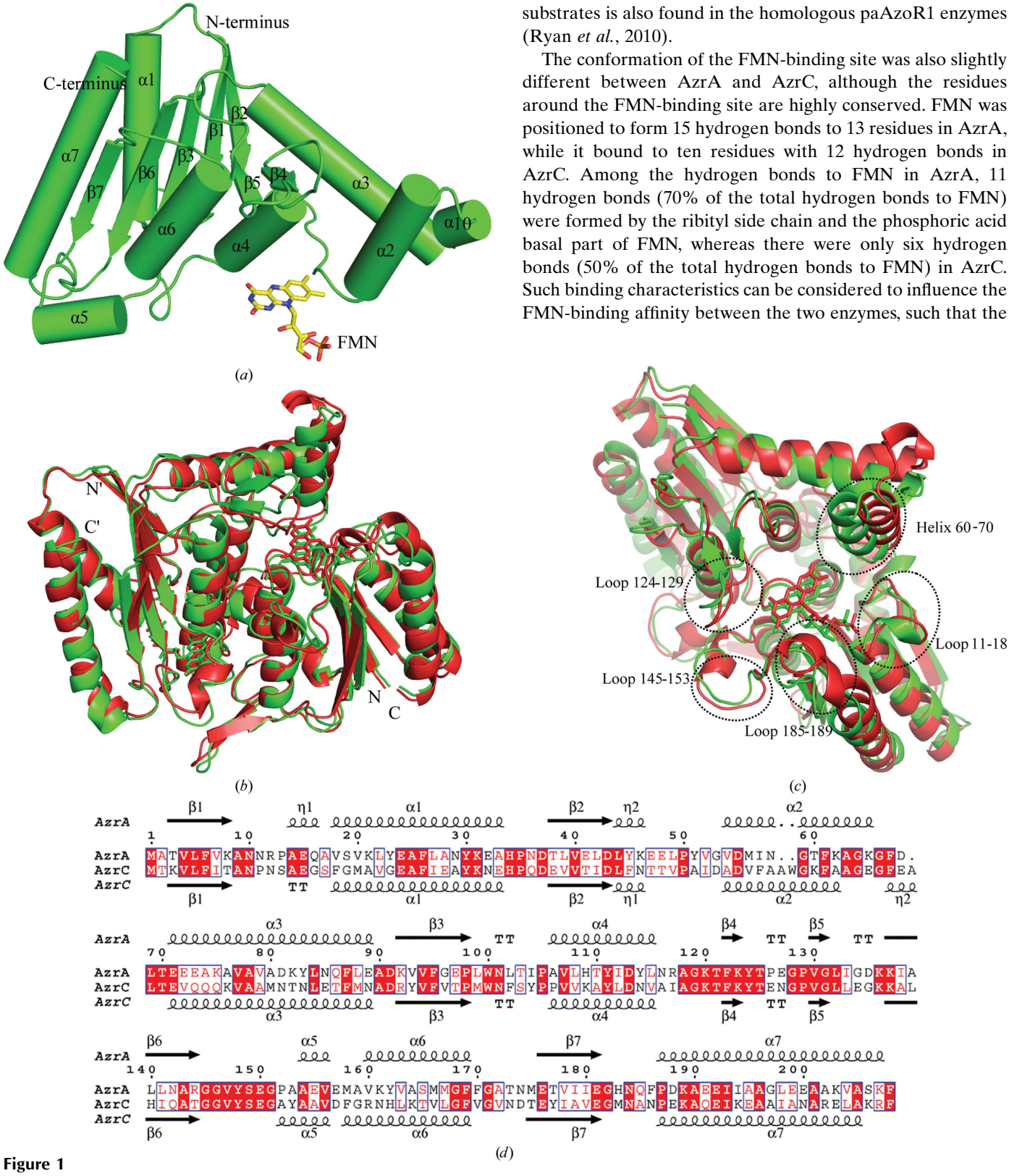


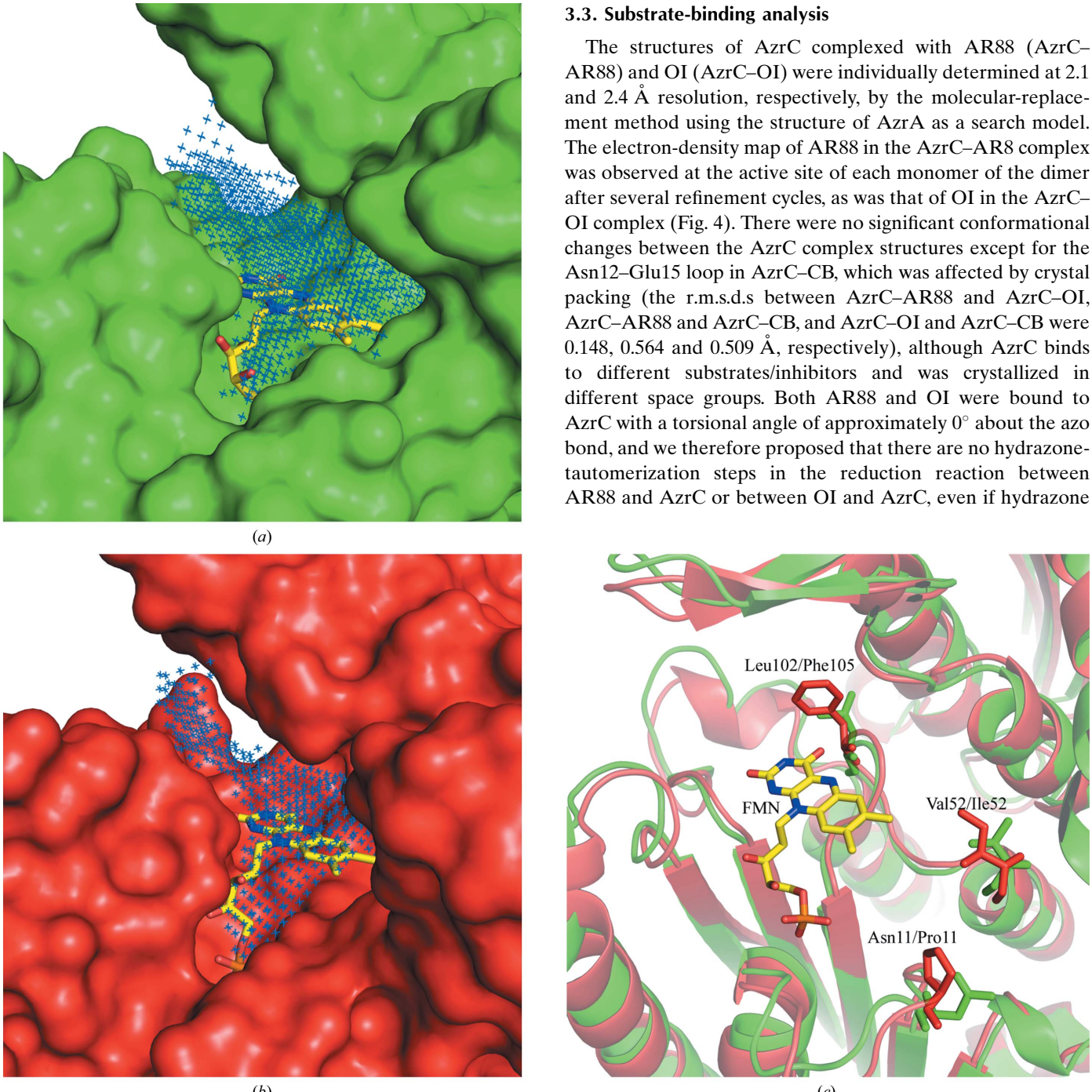
Figure 1
Structural comparison of AzrA and AzrC. The structures of AzrA and AzrC are shown as ribbons. (*a*) The overall structure of the AzrA–FMN complex. FMN (yellow) is shown as sticks. (*b*) Overall structure comparison. AzrA is shown in green and AzrC is shown in red. (*c*) The regions with different conformations around the active site in AzrA and AzrC are circled with dashed lines. (*d*) Sequence alignment between AzrA and AzrC. Secondary structures are also shown above the sequences. All figures in this work were generated using *PyMOL* (Schrödinger).

dissociation constant, K_d, of FMN in AzrA (2.8 n*M*; Ooi *et al.*, 2012) is about half that in AzrC (6.1 n*M*; unpublished results).

As CB is an inhibitor/analogue of NAD(P)H, we solved the structure of the AzrC–CB complex and indirectly provided structural evidence for NAD(P)H binding. CB in the complex was located on top of FMN, similar to the position of substrate in the active site (Figs. 3*b* and 3*c*). The triple ring of CB, which mimics the nicotinamide adenine of NAD(P)H, was sandwiched by Tyr127 and FMN, in the opposite direction to the isoalloxazine ring of FMN. This localization of CB allowed us to build a binding model of NAD(P)H and has implications for the binding position of the nicotinamide adenine of NAD(P)H, which enables C4N of NAD(P)H to transfer the hydride ion to N5 of FMN, which is within a distance of ∼3.5 Å (Fig. 3*d*). The arrangement of CB also suggested that NAD(P)H may bind to Azr in a similar manner.

3.3. Substrate-binding analysis

The structures of AzrC complexed with AR88 (AzrC–AR88) and OI (AzrC–OI) were individually determined at 2.1 and 2.4 Å resolution, respectively, by the molecular-replacement method using the structure of AzrA as a search model. The electron-density map of AR88 in the AzrC–AR8 complex was observed at the active site of each monomer of the dimer after several refinement cycles, as was that of OI in the AzrC–OI complex (Fig. 4). There were no significant conformational changes between the AzrC complex structures except for the Asn12–Glu15 loop in AzrC–CB, which was affected by crystal packing (the r.m.s.d.s between AzrC–AR88 and AzrC–OI, AzrC–AR88 and AzrC–CB, and AzrC–OI and AzrC–CB were 0.148, 0.564 and 0.509 Å, respectively), although AzrC binds to different substrates/inhibitors and was crystallized in different space groups. Both AR88 and OI were bound to AzrC with a torsional angle of approximately 0° about the azo bond, and we therefore proposed that there are no hydrazone-tautomerization steps in the reduction reaction between AR88 and AzrC or between OI and AzrC, even if hydrazone



(*a*) (*b*) (*c*)

Figure 2
Active-site comparison of AzrA and AzrC. AzrA (*a*) and AzrC (*b*) are shown as green and red surfaces, respectively. The calculated binding pockets in AzrA and AzrC are represented by blue points. FMN is shown as sticks; the C atoms of FMN are shown in yellow. (*c*) AzrA and AzrC are shown as green and red ribbons, respectively. The side chains of Asn11/Pro11, Val52/Ile52 and Leu102/Phe105 are shown as green and red sticks, respectively.

tautomerization may facilitate hydride transfer to the azo bond (Ryan *et al.*, 2010).

Interestingly, both AR88 and OI were sandwiched by Tyr127 and FMN in their complexes, but with different manners of binding. While the sulfonate of AR88 was located above the phosphate group of FMN, surrounded by $\alpha2'$ of another molecule, the loop between $\beta1$ and $\alpha1$, and the loop connected to the last helix $\alpha7$ (Fig. 4*a*), the sulfonate of OI was located around the turn before $\beta5'$ of another molecule, the loop after $\beta6$ and the loop connected to $\alpha7$ in a distinct binding orientation to that of AR88 (Fig. 4*b*). Such different locations of AR88 and OI suggested that substrates could bind to AzrC in different ways.

The electron-density maps clearly showed that the azo bond of AR88 was located between Tyr127 OH (N1—OH, 3.87 Å; N2—OH, 3.32 Å) and N10 (N1—N10, 3.50; N2—N10, 3.65 Å) in the middle of the isoalloxazine ring of FMN (the atom labels for FMN, Acid Red 88 and Orange I are as defined

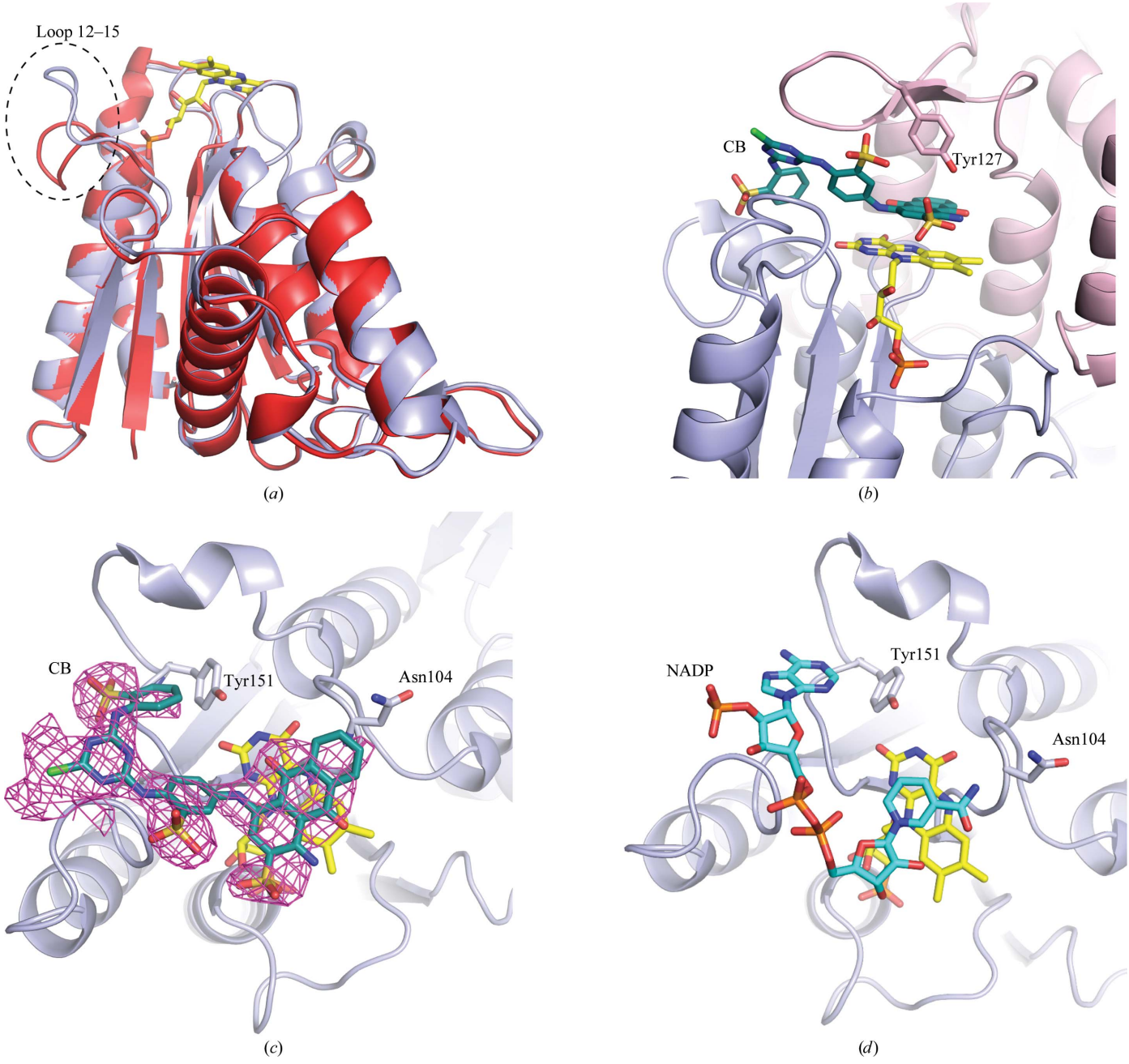


Figure 3
Structure of the AzrC–CB complex. The structure of AzrC is shown as ribbons. FMN, CB and NAD(P)H are shown in stick representation. The C atoms in CB and NAD(P)H are shown in peacock blue and cyan, respectively. CB is bound to AzrC on the top of FMN (yellow) and is sandwiched by Tyr127 and FMN. (*a*) Structure superposition between molecule *A* (light blue) and molecule *B* (red) in the AzrC–CB complex. Loop 12–15 is circled by a dashed line. (*b*) Side view of the AzrC–CB active site. (*c*) Top view of the AzrC–CB active site; the initial $F_o - F_o$ map of CB is contoured at 2σ in purple. (*d*) Proposed binding pose of NAD(P)H.

in Fig. 5). For OI, the azo bond was sandwiched by Pro132 and the pyrimidine ring of FMN (N1—N1, 4.04 Å, N1—N3, 4.05 Å; N2—N1, 3.85 Å, N2—N3, 3.37 Å), whereas the Tyr127 and the middle ring N5–N10 of isoalloxazine sandwiched the naphthalenol (Figs. 4 and 5).

Remarkably, although AR88 and OI were located in the active site in distinctly different orientations, the *ortho*-hydroxy group (*ortho*-HOA1) in AR88 and the *para*-hydroxy group in OI (*para*-HOA1) were located at similar positions and formed hydrogen bonds to the side chain of a conserved Asn (Asn101 in AzrA, Asn104 in AzrC, Asn101 in AzrB and Asn99 in paAzoR1; Fig. 5). Consequently, the *ortho* C atom (CA2) in AR88 and the *para* C atom (CA4) in OI were positioned on the upper site of N5 of FMN at distances of 3.30

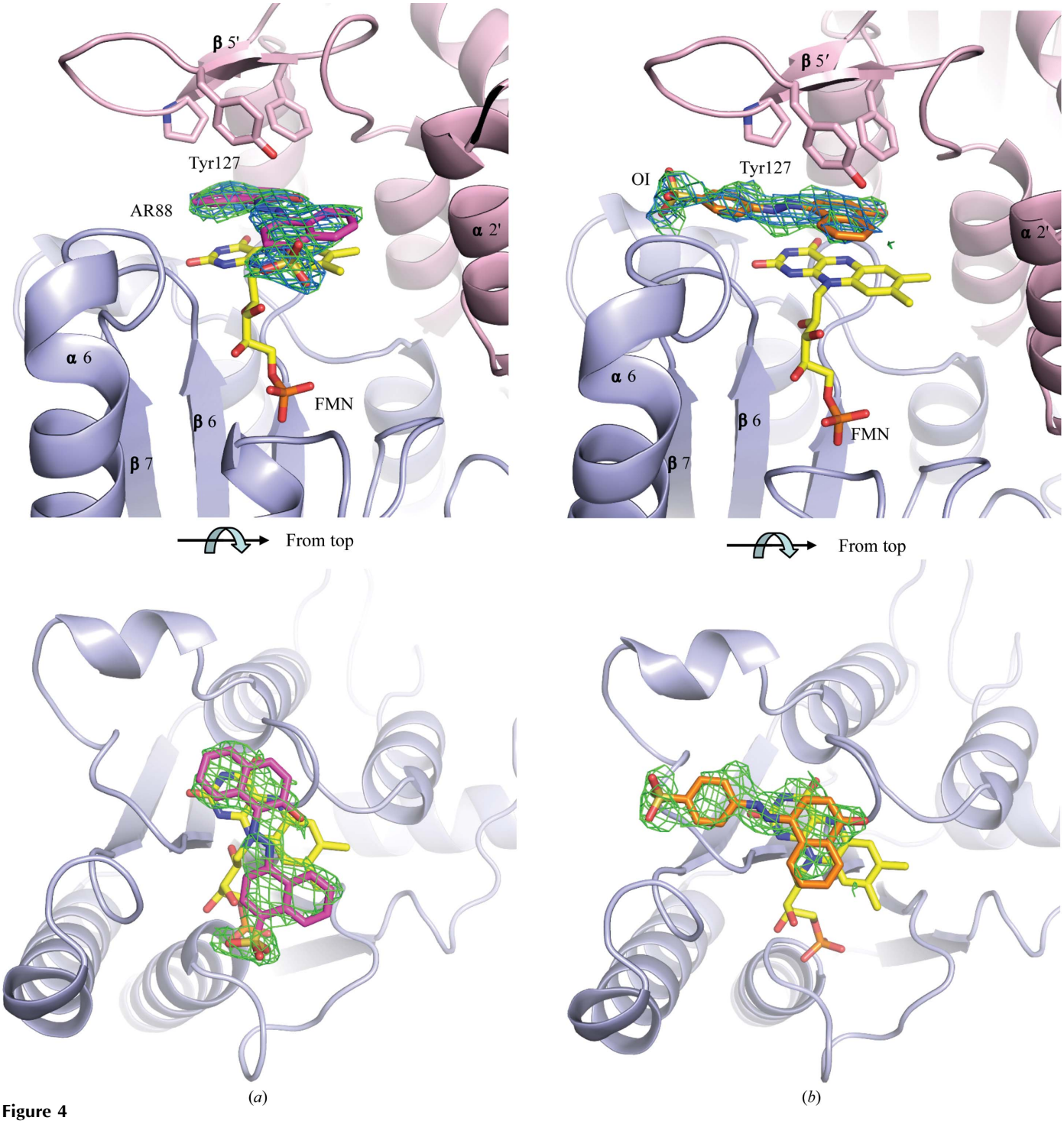


Figure 4
Different substrate-binding poses in the AzrC–AR88 and AzrC–OI complex structures. The substrates and FMN (yellow) are shown as stick representations; the C atoms of AR88 and OI are shown in purple and orange, respectively. The green meshes represent the initial $F_o - F_o$ map contoured at 2σ calculated before substrates were built into the model. (*a*) AzrC–AR88 active site. (*b*) AzrC–OI active site.

and 3.17 Å, respectively. Comparing the residues of the three enzymes around HOA1, we found that most residues were conserved, with only the closest residue being completely different, Ala119 in AzrC and Ser119 in AzrB compared with Asn116 in AzrA, suggesting that the three enzymes AzrA, AzrB and AzrC may possess different catalytic abilities depending on the HOA1 recognition of substrates.

In addition, the binding affinities of AR88 and OI for AzrC were estimated by ITC experiments. The ITC results (Fig. 6) indicated that the binding affinity of AR88 to AzrC (K_a = 2.5 × 10^5 M^{-1}) is very similar to that of OI (K_a = 2.4 × 10^5 M^{-1}).

3.4. Site-directed mutagenesis for reaction mechanisms

It has been reported that human NAD(P)H-quinone oxidoreductase (NQO1), also known as DT-diaphorase, is an azoreductase protein (Wu *et al.*, 1997). In addition, based on a comparison between AzoRs and their structural homologue

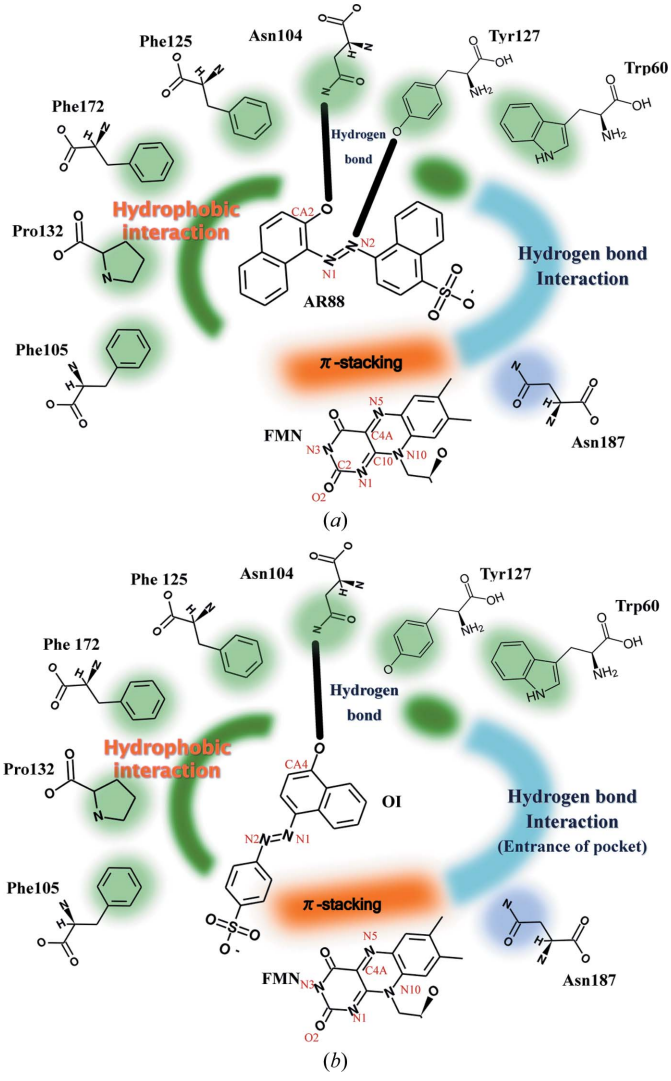


Figure 5
The active sites of AzrC–AR88 (*a*) and AzrC–OI (*b*). The amino-acid residues that surround the substrates around the active sites are labelled. Hydrophobic interactions are shown in green. Hydrogen-bond interactions are shown in blue. π-Stacking interactions are shown in orange.

Table 3
Kinetic data for wild-type and mutant Azrs.

Enzyme	$K_{m,NADH}$ (μ*M*)	$K_{m,\ MR}$ (μ*M*)	$V_{max.\ MR}$ (μ*M* min^{-1})
AzrA	60.8 ± 5.6	79.4 ± 14	41.7 ± 7.1
$AzrA_{Y148A}$	95.0 ± 7.0	119 ± 15	21.7 ± 2.7
AzrC	193 ± 8.1	81.4 ± 3.0	35.3 ± 1.6
$AzrC_{Y151A}$	177 ± 1.9	50.6 ± 8.5	32.8 ± 4.9

NQO1 from *Homo sapiens*, previous studies showed that the reduction mechanism of FMN using NADH bound to azoreductases is similar to that of NQO1 (Ito *et al.*, 2008). Around the N1 and O2 atoms of FMN in NQO1, Gly150, Tyr155 and His161 form a charge-keeper motif and play a key role in the electron-transfer process (Li *et al.*, 1995; Chen *et al.*, 1999). Moreover, it has been suggested that in AzoR from *E. coli*, His141, corresponding to Tyr155 in NQO1, functions as a helper to stabilize the negative charge of the hydride (proton donor) for the reduction reaction (Ito *et al.*, 2008; Ryan *et al.*, 2010).

Sequence comparison around the charge-keeper motif in the Azr family showed that the azoreductase enzymes can be classified into two groups: the His group, in which the His residue is conserved to form a hydrogen bond to O2 of FMN, and the Tyr group, in which the Tyr residue is conserved to form hydrogen bonds to O2 and N1 of FMN. The azo-

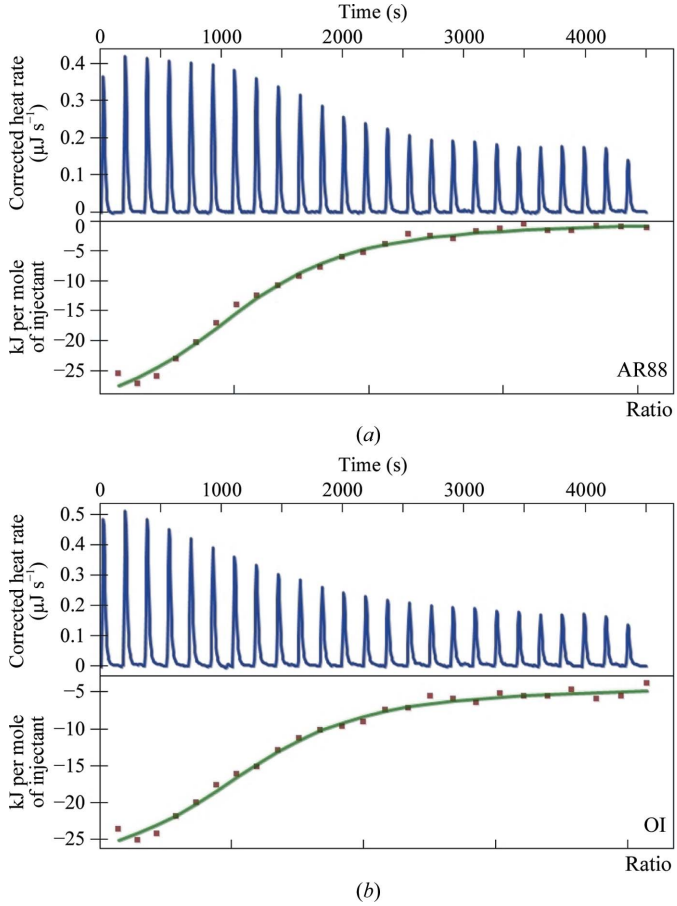


Figure 6
ITC results for AzrC–AR88 (*a*) and AzrC–OI (*b*).

reductases in this study belonged to the Tyr group (Tyr148 in AzrA and AzrB and Tyr151 in AzrC). To confirm the function of Tyr148/151 in Azrs, site-directed mutagenesis and kinetic analysis of $AzrA_{Y148A}$ and $AzrC_{Y151A}$ mutants were carried out (Table 3). The K_m values of the $AzrA_{Y148A}$ and $AzrC_{Y151A}$ mutants for electron transfer from NADH were 0.64-fold and 1.1-fold those of the wild-type AzrA and AzrC, respectively. No significant differences were observed. Therefore, it was suggested that Tyr148 in AzrA and Tyr151 in AzrC are not key residues in the electron-transfer mechanism. Presumably, the main function of Tyr148/151 in Azrs is to stabilize the binding of FMN by hydrogen bonding. Furthermore, these residues also have the potential to subserve other residues involved in the flavin electron-transfer reaction.

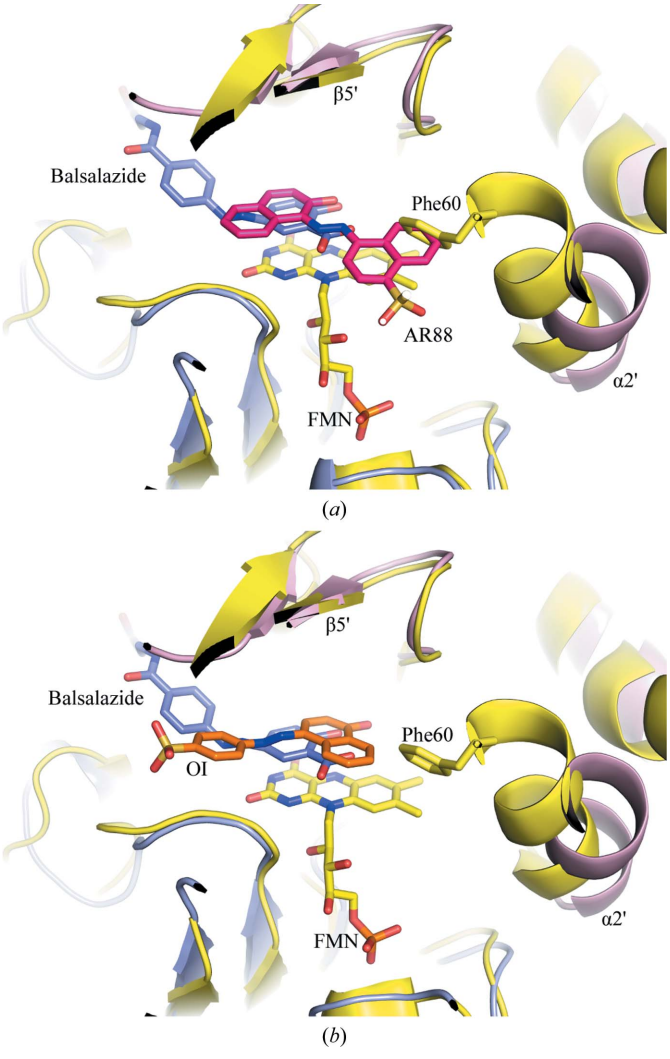


Figure 7
Binding comparison between AR88 and balsalazide (*a*) and between OI and balsalazide (*b*). paAzoR1 (shown as a yellow cartoon) was superposed on AzrC (light pink). The substrates and FMN (yellow) are shown as sticks; the C atoms of balsalazide, AR88 and OI are shown in slate, purple and orange, respectively. AzrC can accommodate both type 1 (AR88) and type 2 (OI) substrates at the active site, while the α2 helix of paAzoR1 prevents the binding of AR88 (a type 2 substrate), as the superposed structures show.

3.5. Substrate specificity

To understand the difference in substrate specificity between AzrA and AzrC, the relative activity ratios of the substrates have been measured (Ooi *et al.*, 2009). It is clear that AzrC has a lower activity than AzrA towards substrates with a naphthalene ring or a sulfone group. On the other hand, for substrates with high hydrophobicity, such as 1-(2-pyridyl-azo)-2-naphthol and Acid Red 88, AzrC shows a tenfold higher activity than that for both Orange I and Orange II (Ooi *et al.*, 2009). The structures of AzrA and AzrC show that AzrC has a higher hydrophobicity and a smaller active-site pocket than AzrA. Therefore, AzrA has a high affinity for larger substrates, while AzrC shows a preference for more hydrophobic substrates. The unconserved assistant residues in the regions of the active-site pocket may play important roles in the substrate specificity of azoreductases.

The structures of AzrC–AR88 and AzrC–OI show different manners of substrate binding, indicating that Azr enzymes can bind substrates in different orientations depending on the characteristics of the substrate. To locate the functional group of substrates above N5 of FMN, substrates with a functional group at the *para* position of the azo bond such as AR88 (type 1) bind to the active site through entrance 1 (formed by $\alpha2'$ of another molecule, the loop between $\beta1$ and $\alpha1$, and the loop connected to the last helix $\alpha7$), while substrates with a functional group at the *ortho* position, such as OI (type 2), should access the active-site pocket from entrance 2 (the turn before $\beta5'$ of another molecule, the loop after $\beta6$ and the loop connected to $\alpha7$). In AzrA, there are more hydrophilic residues around entrance 2 than in AzrC. Such hydrophilic environments are also reflected by the strong affinity and substrate specificity that AzrA has for substrates (Orange I, Orange II and Acid Red 88) with a hydrophilic functional group and a sulfone group.

Compared with another azoreductase, *P. aeruginosa* azoreductase (paAzoR1; Ryan *et al.*, 2010), we found that AzrC may have a broader substrate specificity since the $\alpha2$ helix in AzrC stays relatively far away from the active site, which allows the binding of larger azo-bond substrates, while the $\alpha2$ helix of paAzoR1 seems to intensively resist large substrates of type 1 binding at the active site. The superposed structures of paAzroR1 and AzrC are shown in Fig. 7. It is found that there is a severe clash between the naphthalenol moiety next to the sulfone group in AR88 and Phe60 of the $\alpha2$ helix in paAzoR1.

3.6. Reduction reaction mechanism

The enzymatic reaction of azoreductases has been reported to involve the cleavage of azo bonds catalyzed by a ping-pong mechanism to reduce the azo compound (Nakanishi *et al.*, 2001; Ooi *et al.*, 2007, 2009). FMN in the active-site pocket can be reduced to the reduced form by NADH. Generally, oxidized FMN can be reduced by a single electron to semiquinone FMN, which is protonated at N5 and pools the negative charge in the area of C4A, C10, N1, C2 and O2 of the isoalloxazine ring. The semiquinone FMN can be reduced by

the addition of a single proton to become reduced FMN protonated at the N1 or O2 position (Cuello *et al.*, 2000; Miura, 2001; Bornemann, 2002; Macheroux *et al.*, 2005). Similar to AzoRs, EmoB and NQO1, AzrA and AzrC have no residue that is able to donate a proton to N1, while there are two residues (Gly146/149 and Tyr148/151) that can donate a proton to the O2 position. The results of the present experiments showed that type 1 family azoreductases can be classified into two subgroups according to the charge-keeper motif: one conserves the Gly-Tyr pair, as represented by AzrA and AzrC, while the other group has a Gly-His pair, as typified by AzoR. The proposed mechanism for the reduction of FMN by NADH in Azrs is similar to that in AzoRs and EmoB (Nissen *et al.*, 2008). The two residues Gly145/Gly148 and Gly146/Glu149 (in AzrA/AzrC, respectively) function as an electron keeper, while Tyr148/Tyr151 only helps to stabilize the binding of FMN.

Generally, NAD(P)H must be released from the enzyme before the substrate can bind. The reduced FMN can provide one or two hydrides from N atoms N1 or N5 by one or two cycles in Azrs. Both the AzrC–AR88 and AzrC–OI complex structures show that the azo bond of the substrate is distant from the N5 atom of FMN, but the electrophilic C atom (CA2 or CA4) that binds to a hydroxy group is positioned within proton/electron-donor distance (∼3.5 Å). We also compared our structures with the paAzoR1–balsalazide complex structure (PDB entry 3lt5; Ryan *et al.*, 2010) and found that the two *para*-position C atoms (labelled CAD) in balsalazides, each of which is also an electrophilic C atom, bound to a hydroxy group close to N5 of FMN, with distances of 3.51 and 3.54 Å, respectively.

4. Conclusion

Based on the present structures, we propose that substrates can be recognized using the following mechanism in azoreductases. Firstly, the substrates enter the active site of the azoreductase in one of two ways depending on the position of the functional group in the naphthalenol part. The hydrophobic parts of the substrate (if they exist) then interact with several hydrophobic residues in the substrate-binding pocket, and the hydrophilic group of the substrate is located at a special position and fixed by hydrogen bonding to a conserved asparagine residue (Asn101 in AzrA and Asn104 in AzrC). Finally, the azo bond is sandwiched and stabilized by π-stacking with Tyr127 and FMN. Azrs could not bind tightly to substrates such as Ethyl Red, Allura Red and Orange (Ooi *et al.*, 2009) as they possess a larger functional group than COOH in the naphthalenol part.

Taken together, we propose two types of attack for the reduction reaction depending on the type of substrate: (i) an electrophilic C atom with a hydroxy group at the *ortho* position (CA2 in AR88) and (ii) an electrophilic C atom with a hydroxy group at the *para* position (CA4 in OI and CAD in

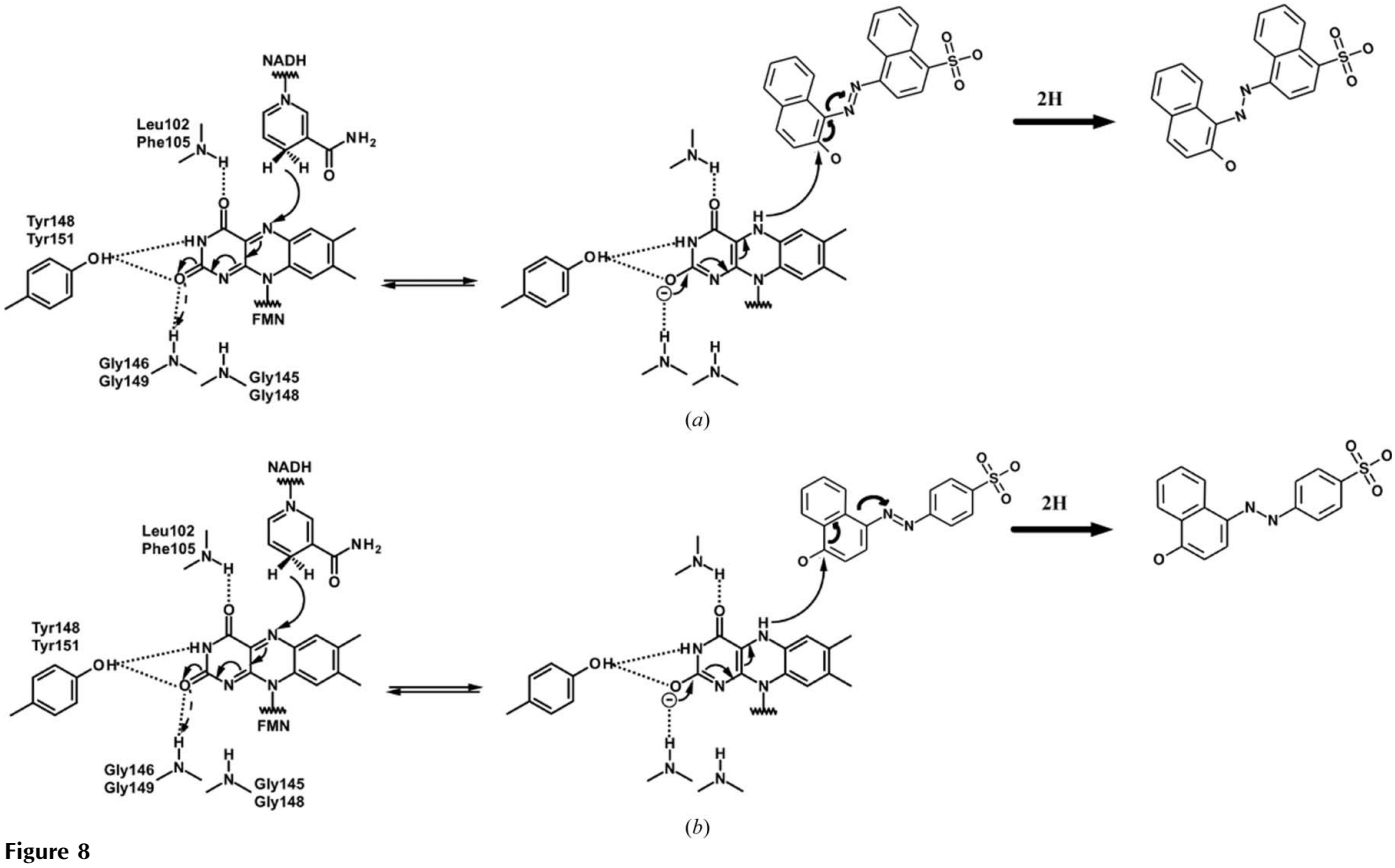


Figure 8
Catalytic reaction mechanism for the reduction of (*a*) AR88 and (*b*) OI.

balsalazide). For both of the two types the substrate accepts hydrides indirectly from N5 of reduced FMN and transfers them to the azo bond (Fig. 8). The reduction process is probably as follows: the azo bond N═N is reduced by reduced $FMNH_2$ to a single HN—NH bond; the product with the HN—NH bond is then released and finally the HN—NH bond accepts a proton from bulk solvent and cleavage is performed without the enzyme. This final step was considered to be the first step of the reduction mechanism since hydrazone tautomerization was supported by the bent bound conformation of balsalazide (Ryan *et al.*, 2010), and the sp^3 hybridization may facilitate balsalazide accepting a proton from the solvent. However, AR88 and OI bound to AzrC in a relatively stable sp^2 hybridization state, so we propose that the HN—NH bond is cleaved in the final step of the reduction mechanism.

We thank staff of the beamlines at the Photon Factory and SPring-8, Japan for their help with data collection. Min Yao is supported by National Natural Science Foundation of China (Grant No. 31370731).

References

Bin, Y., Jiti, Z., Jing, W., Cuihong, D., Hongman, H., Zhiyong, S. & Yongming, B. (2004). *FEMS Microbiol. Lett.* **236**, 129–136.

Binter, A., Staunig, N., Jelesarov, I., Lohner, K., Palfey, B. A., Deller, S., Gruber, K. & Macheroux, P. (2009). *FEBS J.* **276**, 5263–5274.

Blümel, S., Knackmuss, H. J. & Stolz, A. (2002). *Appl. Environ. Microbiol.* **68**, 3948–3955.

Bornemann, S. (2002). *Nat. Prod. Rep.* **19**, 761–772.

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Brunger, A. T. (2007). *Nature Protoc.* **2**, 2728–2733.

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

Chen, H. (2006). *Curr. Protein Pept. Sci.* **7**, 101–111.

Chen, H., Hopper, S. L. & Cerniglia, C. E. (2005). *Microbiology*, **151**, 1433–1441.

Chen, H., Wang, R. F. & Cerniglia, C. E. (2004). *Protein Expr. Purif.* **34**, 302–310.

Chen, S., Wu, K., Zhang, D., Sherman, M., Knox, R. & Yang, C. S. (1999). *Mol. Pharmacol.* **56**, 272–278.

Chung, K.-T. (1983). *Mutat. Res.* **114**, 269–281.

Chung, K.-T., Fulk, G. E. & Andrews, A. W. (1981). *Appl. Environ. Microbiol.* **42**, 641–648.

Cuello, A. O., McIntosh, C. M. & Rotello, V. M. (2000). *J. Am. Chem. Soc.* **122**, 3517–3521.

Dissanayake, A. S. & Truelove, S. C. (1973). *Gut*, **14**, 923–926.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Griffiths, J. (1984). *Developments in the Chemistry and Technology of Organic Dyes*, edited by J. Griffiths, p. 30. Oxford: Blackwell.

Hanauer, S. B. (1996). *N. Engl. J. Med.* **334**, 841–848.

Ito, K., Nakanishi, M., Lee, W.-C., Sasaki, H., Zenno, S., Saigo, K., Kitade, Y. & Tanokura, M. (2005). *Acta Cryst.* F**61**, 399–402.

Ito, K., Nakanishi, M., Lee, W.-C., Sasaki, H., Zenno, S., Saigo, K., Kitade, Y. & Tanokura, M. (2006). *J. Biol. Chem.* **281**, 20567–20576.

Ito, K., Nakanishi, M., Lee, W.-C., Zhi, Y., Sasaki, H., Zenno, S., Saigo, K., Kitade, Y. & Tanokura, M. (2008). *J. Biol. Chem.* **283**, 13889–13896.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kariminiaae-Hamedaani, H.-R., Sakurai, A. & Sakakibara, M. (2007). *Dyes Pigm.* **72**, 157–162.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Li, R., Bianchet, M. A., Talalay, P. & Amzel, L. M. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 8846–8850.

Lineweaver, H. & Burk, D. (1934). *J. Am. Chem. Soc.* **56**, 658–666.

Liu, Z.-J., Chen, H., Shaw, N., Hopper, S. L., Chen, L., Chen, S., Cerniglia, C. E. & Wang, B.-C. (2007). *Arch. Biochem. Biophys.* **463**, 68–77.

Liu, G., Zhou, J., Fu, Q. S. & Wang, J. (2009). *J. Bacteriol.* **191**, 6394–6400.

Macheroux, P., Ghisla, S., Sanner, C., Rüterjans, H. & Müller, F. (2005). *BMC Biochem.* **6**, 26.

Matsumoto, K., Mukai, Y., Ogata, D., Shozui, F., Nduko, J. M., Taguchi, S. & Ooi, T. (2010). *Appl. Microbiol. Biotechnol.* **86**, 1431–1438.

Miura, R. (2001). *Chem. Rec.* **1**, 183–194.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nakanishi, M., Yatome, C., Ishida, N. & Kitade, Y. (2001). *J. Biol. Chem.* **276**, 46394–46399.

Nishiya, Y. & Yamamoto, Y. (2007). *Biosci. Biotechnol. Biochem.* **71**, 611–614.

Nissen, M. S., Youn, B., Knowles, B. D., Ballinger, J. W., Jun, S.-Y., Belchik, S. M., Xun, L. & Kang, C. (2008). *J. Biol. Chem.* **283**, 28710–28720.

Ogata, D., Ooi, T., Fujiwara, T., Taguchi, S., Tanaka, I. & Yao, M. (2010). *Acta Cryst.* F**66**, 503–505.

Ooi, T., Ogata, D., Matsumoto, K., Nakamura, G., Yu, J., Yao, M., Kitamura, M. & Taguchi, S. (2012). *J. Mol. Catal. B Enzym.* **74**, 204–208.

Ooi, T., Shibata, T., Matsumoto, K., Kinoshita, S. & Taguchi, S. (2009). *Biosci. Biotechnol. Biochem.* **73**, 1209–1211.

Ooi, T., Shibata, T., Sato, R., Ohno, H., Kinoshita, S., Thuoc, T. L. & Taguchi, S. (2007). *Appl. Microbiol. Biotechnol.* **75**, 377–386.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.

Pedraza, A., Sicilia, M. D., Rubio, S. & Pérez-Bendito, D. (2006). *Analyst*, **131**, 81–89.

Ryan, A., Laurieri, N., Westwood, I., Wang, C.-J., Lowe, E. & Sim, E. (2010). *J. Mol. Biol.* **400**, 24–37.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Stolz, A. (2001). *Appl. Microbiol. Biotechnol.* **56**, 69–80.

Sugiura, W., Yoda, T., Matsuba, T., Tanaka, Y. & Suzuki, Y. (2006). *Biosci. Biotechnol. Biochem.* **70**, 1655–1665.

Suzuki, Y., Yoda, T., Ruhul, A. & Sugiura, W. (2001). *J. Biol. Chem.* **276**, 9059–9065.

Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 38–44.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.

Wang, C.-J., Hagemeier, C., Rahman, N., Lowe, E., Noble, M., Coughtrie, M., Sim, E. & Westwood, I. (2007). *J. Mol. Biol.* **373**, 1213–1228.

Wu, K., Knox, R., Sun, X. Z., Joseph, P., Jaiswal, A. K., Zhang, D., Deng, P. S.-K. & Chen, S. (1997). *Arch. Biochem. Biophys.* **347**, 221–228.

Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.

Yu, J., Zhou, Y., Tanaka, I. & Yao, M. (2010). *Bioinformatics*, **26**, 46–52.