Crystal structure of thebaine 6-O-demethylase from the morphine biosynthesis pathway

Anna Kluza[1], Ewa Niedzialkowska[1,a], Katarzyna Kurpiewska[1,2], Zuzanna Wojdyla[1], Matthew Quesne[1,b], Ewa Kot[1], Przemyslaw J. Porebski[1,a]*, Tomasz Borowski[1]*

[1]Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30239 Krakow, Poland

[2]Department of Crystal Chemistry and Crystal Physics, Faculty of Chemistry, Jagiellonian University, 30387 Krakow, Poland

Present address:

[a]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908, USA

[b]School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff, CF10 3AT, UK

AUTHOR INFORMATION

Contact details to corresponding authors:

*Przemyslaw J. Porebski:

pjp2b@virginia.edu, +1 434-243-0033

Department of Molecular Physiology and Biological Physics

University of Virginia

1340 Jefferson Park Avenue, Pinn Hall, Charlottesville, VA 22908, USA

*Tomasz Borowski:

ncborows@cyf-kr.edu.pl, +48 12-639-5158

Jerzy Haber Institute of Catalysis and Surface Chemistry

Polish Academy of Sciences

Niezapominajek 8, PL-30239 Krakow, Poland

KEYWORDS

thebaine 6-O-demethylase, T6ODM, morphine biosynthesis, O-demethylation, X-ray crystallography, 2-oxoglutarate/Fe(II)-dependent dioxygenases

ABBREVIATIONS

2OG, 2-oxoglutarate; ANS, anthocyanidin synthase; CODM, codeine 3-O-demethylase; F6'H, feruloyl-CoA 6'-hydroxylase; ODD, 2-oxoglutarate/Fe(II)-dependent dioxygenases; T6ODM, thebaine 6-O-demethylase; SIN, succinate.

ABSTRACT

Thebaine 6-O-demethylase (T6ODM) from *Papaver somniferum* (opium poppy) is a key enzyme in the morphine biosynthesis pathway that belongs to the non-heme 2-oxoglutarate/Fe(II)-dependent dioxygenases (ODD) family. Initially, T6ODM was characterized as an enzyme catalyzing O-demethylation of thebaine to neopinone and oripavine to morphinone, however recently the substrate range of T6ODM was expanded to a number of various benzylisoquinoline alkaloids. Here, we present crystal structures of T6ODM in complexes with 2-oxoglutarate (T6ODM:2OG, PDB: 5O9W) and succinate (T6ODM:SIN, PDB: 5O7Y). The arrangement of the T6ODM's active site is typical for proteins from the ODD family, but the enzyme is characterized by a large substrate binding cavity, whose volume can partially explain the T6ODM promiscuity. Moreover, the size of the cavity allows for binding of multiple molecules at once, posing a question about substrate-driven specificity of the enzyme.

T6ODM from opium poppy

Non-heme 2-oxoglutarate/Fe(II)-dependent dioxygenases are widely distributed in nature and catalyze a wealth of modifications, e.g. hydroxylation, demethylation, desaturation, epimerization. Recently, a group of new ODD has been identified that encompasses enzymes involved in biosynthesis of morphinan alkaloids in opium poppy. Thebaine 6-O-demethylase (T6ODM) and codeine-3-O-demethylase (CODM) were long supposed to belong to cytochrome P450 family, until 2010 when it was revealed that these closely related enzymes belong to an ODD family. Notably, T6ODM and CODM are the first members of this family demonstrated to carry out O-demethylation (Hagel and Facchini, 2010).

The final stage of morphine biosynthesis starts from thebaine and continues by two routes, both of which engage codeinone reductase (COR) and two ODD enzymes: T6ODM and CODM. T6ODM catalyzes the removal of the methyl group from the C6-bound oxygen atom of thebaine and oripavine, whereas CODM removes the methyl group from the C3-bound oxygen atom of thebaine and codeine (**Figure 1**). However, the substrate range of T6ODM and CODM has recently been studied and proven to be broader than initially expected. It encompasses different benzylisoquinoline alkaloids from the protopine, protoberberine, aporphine and 1-benzylisoquinoline groups (Farrow and Facchini, 2013).

The fact that T6ODM and CODM belong to the same enzyme family, with 73% identity and 85% similarity at the amino acid sequence level, raises the question about the origins of the regioselectivity toward the demethylation reactions. Previously, this question was addressed by W. Runguphan and colleagues who attempted to swap T6ODM and CODM reaction specificities (Runguphan et al., 2012). They removed the demethylase activity toward thebaine from CODM by substituting four amino acids (E338G, I340L, L341V, K342E), which are relatively remote from the active site. However, these mutations did not manage to completely change the specificity. Due to lack of the crystal structure of CODM or T6ODM, the structural model used in that study was based on the closest homolog (30% of identity at amino acid sequence level) with a known 3D structure, anthocyanidin synthase (ANS) from *Arabidopsis thaliana* (Wilmouth et al., 2002).

Here we present the crystal structures of T6ODM in complexes with 2OG (PDB: 5O9W) and SIN (PDB: 5O7Y) that may serve as a starting point for further structural and mutagenesis studies and, hopefully, will help in the understanding of the origins of regioselectivity during the catalytic reactions of T6ODM and CODM.

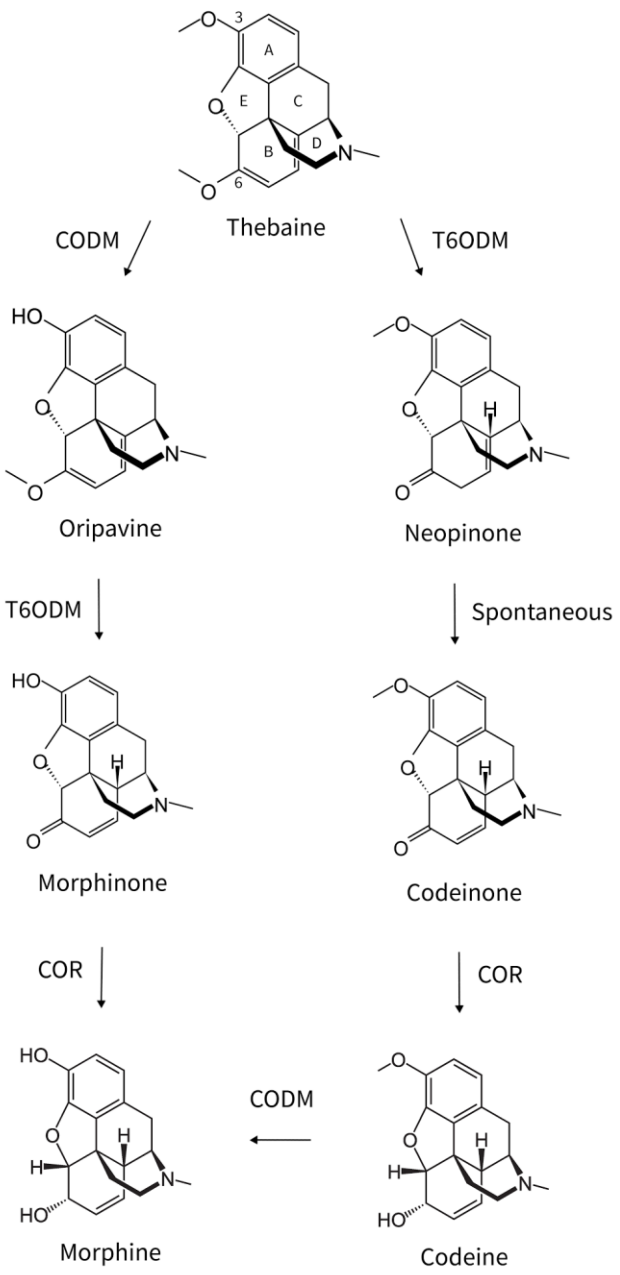

Figure 1. Final stage of morphine biosynthesis pathway starts from thebaine and continues by two routes. T6ODM and CODM take part in both of them. Despite high similarity at amino acid sequence level they differ in reaction regioselectivity. T6ODM demethylates thebaine and oripavine at C6-bound methoxy group, while CODM demethylates thebaine and codeine at C3-bound methoxy group. Codeinone and

morphinone are reduced by COR to codeine and morphine, respectively.

Protein expression and purification

A codon-optimized gene encoding the T6ODM protein (Uniprot: D4N500) was synthesized by GenScript (Piscataway, NJ) and cloned into a modified pET28 vector. The construct was designed to allow direct expression from pET28 vector, subcloning into MCSG vectors using a ligation independent cloning (Eschenfeldt et al., 2009) or into StarGate system vectors (IBAGmbH, Germany). The expression was tested using pET28, pMCSG7 and pMCSG19 vectors and *Escherichia coli* strains BL21(DE3)Ripl, BL21(DE3)Arctic and BL21(DE3)Magic with best expression and solubility levels resulting from pMCSG19_T6ODM construct and BL21(DE3)Magic cells. As a result, T6ODM was overproduced as $(His)_6$-MBP-T6ODM fusion protein with a cleavage site for Tobacco Etch Virus (TEV) protease.

E. coli strain BL21(DE3)Magic cells harboring pMCSG19_T6ODM construct were grown in 1 L flasks of Terrific Broth medium at 37ºC until optical density at 600 nm reached approx. 1. The overproduction of $(His)_6$-MBP-T6ODM fusion protein was induced with 0.15 mM isopropyl-β-D-1-thiogalactopyranoside and performed for 20 hours at 16ºC with shaking. The cells were harvested by centrifugation at 4000 rcf for 40 min at 4ºC. For purification of $(His)_6$-MBP-T6ODM fusion protein, bacterial pellets were resuspended (Lysis Buffer: 50 mM Tris-HCl, 150 mM NaCl, pH 7.8, 5% (v/v) glycerol) and incubated with DNase (Roche) and lysozyme (Carl Roth). The cell suspension was lysed by sonication on ice in the presence of a EDTA-free Protease Inhibitor Coctail (Roche). After clarification by centrifugation for 1 hour at 40000 rcf, the supernatant was applied to a Ni-NTA Superflow resin (Qiagen). Unbound proteins were washed away (Washing Buffer: 50 mM Tris, 600 mM NaCl, 10 mM imidazole, pH 7.8) and the protein of interest eluted (Elution Buffer: 50 mM Tris, 150 mM NaCl, 250 mM imidazole, pH 7.8). Cleavage of $(His)_6$-MBP-T6ODM fusion protein was performed overnight with recombinant $(His)_6$-TEV protease in buffer containing 50 mM Tris, 150 mM NaCl at pH 7.8. $(His)_6$-MBP, $(His)_6$-TEV protease and undigested $(His)_6$-MBP-T6ODM protein were removed by Ni^{2+} affinity chromatography. The NiNTA resin flowthrough containing T6ODM protein was further purified using Superdex200 16/600 column attached to ÄKTApurifier system (Running Buffer: 50 mM Tris, 500 mM NaCl, pH 7.8). Before crystallization, buffer was exchanged with water by diafiltration, and the protein was concentrated to 37.5 mg/ml using Amicon centrifugal filter units (Millipore) with 10 kDa molecular weight cut-off.

Crystallization

Crystals of T6ODM were grown by sitting drop vapour diffusion method at 4ºC using 2.4 µl drops containing equal volumes of protein solution and mother liquor. Preliminary crystallization conditions were

examined using PACT premier (Molecular Dimensions), Wizard Cryo (Rigaku Reagents), SaltRx, Index (Hampton Research) commercial screens. After a series of optimization, microcrystals were grown from modified PACT premier (0.2 M potassium sodium tartrate tetrahydrate, 20% PEG 3350) and SaltRx (0.6 M potassium sodium tartrate tetrahydrate, 0.1 M Tris pH 8.5; 0.6 M potassium sodium tartrate tetrahydrate, 0.1 M BIS-TRIS propane pH 7.0) conditions. Based on the assumption that tartrate may substitute for either the substrate or product of T6ODM, it was replaced with 2OG and SIN in the crystallization buffer. The concentration of T6ODM, PEG 3350, 2OG, SIN as well as the temperature of growth were systematically varied. Finally, the T6ODM:2OG complex was crystallized using 19% PEG 3350 and 200 mM disodium 2-oxoglutarate, pH 9.4 solution as a mother liquor, while the T6ODM:SIN complex crystallized in 22% PEG 3350 and 125 mM disodium succinate, pH 7.9. Plate-shaped crystals of T6ODM:SIN and thick needles of T6ODM:2OG were observed within 3-5 weeks (**Figure S1**). For data collection, crystals were harvested using CryoLoops (Hampton Research), transferred into cryoprotectant solution (mother liquor containing 30% ethylene glycol) and flash-cooled in LN2 cryo-stream (130K).

Crystallization trials of T6ODM with available product analog (morphine) were carried out. However, although crystals did grow in the solution containing morphine, the electron density indicating morphine presence in the active site could not be found.

Data collection and data analysis

The diffraction data was collected at a temperature of 130 K with Rigaku Oxford Diffraction SuperNova dual source diffractometer with a Cu-K_α radiation source and 135 mm Atlas 2 CCD area detector at Faculty of Chemistry, Jagiellonian University. Data were collected using CrysAlis^{Pro} (Oxford Diffraction, 2006) and processed with HKL-2000 (Otwinowski and Minor, 1997). Structure solution and model building were carried out with HKL-3000 (Minor et al., 2006) coupled with MOLREP (Vagin and Teplyakov, 1997), BUCCANEER (Cowtan, 2008) and Fitmunk (Porebski et al., 2016). The T6ODM:SIN structure was solved by molecular replacement with structure of feruloyl-CoA 6-hydroxylase (F6'H) from *Arabidopsis thaliana* as a search model (PDB: 4XAE) (Sun et al., 2015), while T6ODM:2OG was based on T6ODM:SIN model. Obtained models were further refined with REFMAC 5.8 (Murshudov et al., 2011) and rebuilt with COOT 0.8.6 (Emsley et al., 2010). The quality of models was assessed using Molprobity (Chen et al., 2010) and wwPDB Validation Service (Young et al., 2017). Diffraction data have been deposited to The Integrated Resource for Reproducibility in Macromolecular Crystallography (IRRMC, proteindiffraction.org) (Grabowski et al., 2016) and can be accessed at DOI:10.18430/M35O7Y (T6ODM:SIN) and DOI:10.18430/M35O9W (T6ODM:2OG). Identification of the sodium ion in T6ODM:2OG electron density map was based on its environment (Zheng et al., 2017) and validated using CheckMyMetal (CMM) (Zheng et al., 2013). The distances between sodium and

coordinated atoms were restrained to 2.4 Å using REFMAC restraints generated by CMM. The distances and geometry between the nickel and coordinated atoms were unrestrained.

Anomalous data collection and analysis

The X-ray energy scan in the vicinity of absorption K-edge of nickel (8.305 - 8.370 keV) was performed for T6ODM:2OG crystals at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) (Cianci et al., 2017). Two diffraction data sets were collected at energies of 8.360 keV (1.4831 Å) and 8.320 keV (1.4902 Å) - above and below the nickel absorption K-edge, respectively. Data was processed with XDS (Kabsch, 2010) and CCP4 (Winn et al., 2011) coupled with POINTLESS 1.10, SCALA 3.3 (Evans, 2006, 2011) and TRUNCATE 1.17 (French et al., 1974). The T6ODM:2OG model (5O9W) was used for Fourier synthesis using REFMAC 5.8, followed by ten cycles of maximum likelihood restrained refinement using SAD data directly at the wavelength of 1.4902 Å and 1.4831 Å, accordingly. Diffraction data have been deposited to IRRMC.

Overall crystal structure

The crystal structures of T6ODM in its complexes with 2OG (PDB: 5O9W) and SIN (PDB: 5O7Y) were determined by molecular replacement using 4XAE as a search model and refined to 1.85 Å and 1.97 Å, respectively. T6ODM crystallized in the $P2_12_12_1$ space group with one monomer in the asymmetric unit and estimated water content of 48%. Data collection, structure refinement and validation statistics are summarized in **Table 1**. The residues 37-40, 46-52 for T6ODM:2OG and residues 38-40, 46-51 for T6ODM:SIN were not included in the final models due to insufficient electron density.

Despite relatively low amino acid sequence identity, the general backbone of T6ODM shows high similarities to its closest homologs with known 3D structures. ANS (PDB: 1GP4; amino acid sequence identity: 30%) and F6'H (PDB: 4XAE; amino acid sequence identity: 32%) from *Arabidopsis thaliana* can be superposed with T6ODM:2OG with a root mean square deviation (RMSD) of 1.77 Å and 1.62 Å, respectively (**Figure S2**) (Krissinel and Henrick, 2004).

Table 1. Diffraction data collection and refinement statistics. The coordinate data files and structure factors have been deposited in the Protein Data Bank as 5O9W (T6ODM:2OG) and 5O7Y (T6ODM:SIN). Diffraction data have been deposited to IRRMC.

PDB ID	5O9W	5O7Y
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
$\alpha = \beta = \gamma$ [°]	90	90
a, b, c [Å]	47.04, 87.42, 94.86	47.23, 87.03, 94.17
Resolution [Å]	1.85 (1.85 - 1.88)	1.97 (1.97 - 2.00)
R_{merge}	0.070 (0.355)	0.136 (0.653)
I / σ (I)	15.773 (2.271)	17.710 (2.271)
Completeness [%]	99 (95)	100 (100)
Redundancy	3.4 (2.2)	15.8 (7.5)
Refinement		
Resolution range [Å]	41.43 - 1.85	41.51 - 1.97
Completeness for range [%]	97.57	99.42
Number of reflections	31662	26620
R_{work} / R_{free}	0.162 / 0.214	0.163 / 0.207
Number of atoms	3220	3206
B factors [Å^2]	24.61	29.24
Structure quality		
R.m.s deviations		
Bond lengths [Å]	0.010	0.009
Bond angles [°]	1.414	1.387
Ramachandran statistics		
Favoured [%]	97.9	98.86
Allowed [%]	99.7	99.72
Outliers [%]	0.29	0.28
Molprobity analysis		
Clashscore	1.32 (100th percentile)	1.69 (100th percentile)
Poor rotamers [%]	0.00	0.00
Favoured rotamers [%]	96.70	98.12
Molprobity score	0.88 (100th percentile)	0.92 (100th percentile)

The core of T6ODM consists of multiple antiparallel β-strands, eight of which form a β-jellyroll fold conserved throughout the ODD superfamily (**Figure 2**). The active site is sandwiched between two β-sheets, major (strands β5, β7, β10, β12) and minor (strands β6, β8, β9, β11). A highly conserved HXD/E...H motif (responsible for metal binding) is located in strand β6 (residue His-238), loop β6/β7 in the close proximity to β6 (Asp-240) and strand β11 (His-295). Structural studies of ODD have revealed that strands β5, β6 and β12, loops β5/β6, β6/β7, β8/β9 as well as the N-terminal and C-terminal regions play an important role in substrate and co-substrate binding. In some cases, loop β8/β9 can also take part in substrate binding or dimer formation, although, this region varies significantly in size and function among ODD (Aik et al., 2012). In the case of T6ODM, this loop consists of only four amino acids and according to PDBePISA analysis (Krissinel and Henrick, 2007) does not facilitate any contacts that can be responsible for dimerization. The main core of T6ODM is surrounded by additional six β-strands, twelve α-helices and four 3_{10} helices (according to DSSP). The long α8 helix in the N-terminal region is placed in immediate proximity to the β-jellyroll fold, which is consistent with the previously reported ANS and F6'H structures, and might act as a 'structural backbone' (Sun et al., 2015; Wilmouth et al., 2002).

We have observed a possible modification of Lys-3 in T6ODM:2OG structure, but as a crystallization artifact irrelevant to structural studies of T6ODM it has not been further investigated and verified. Similar, but more disordered density was observed for Lys-336. These additional densities were assigned as an unknown atoms (UNX) and can be inspected interactively using Molstack (Porebski et al., 2017) at http://molstack.bioreproducibility.org/project/view/OJM60NCMQF1VMKW0IBUF. Because these additional densities were not present in the T6ODM:SIN structure we suspect that, due to high concentration of 2OG (200 mM), 2OG underwent a non-enzymatic condensation reaction with a lysine

yielding a saccharopine (known as an intermediate in lysine degradation) (Sacksteder et al., 2000).

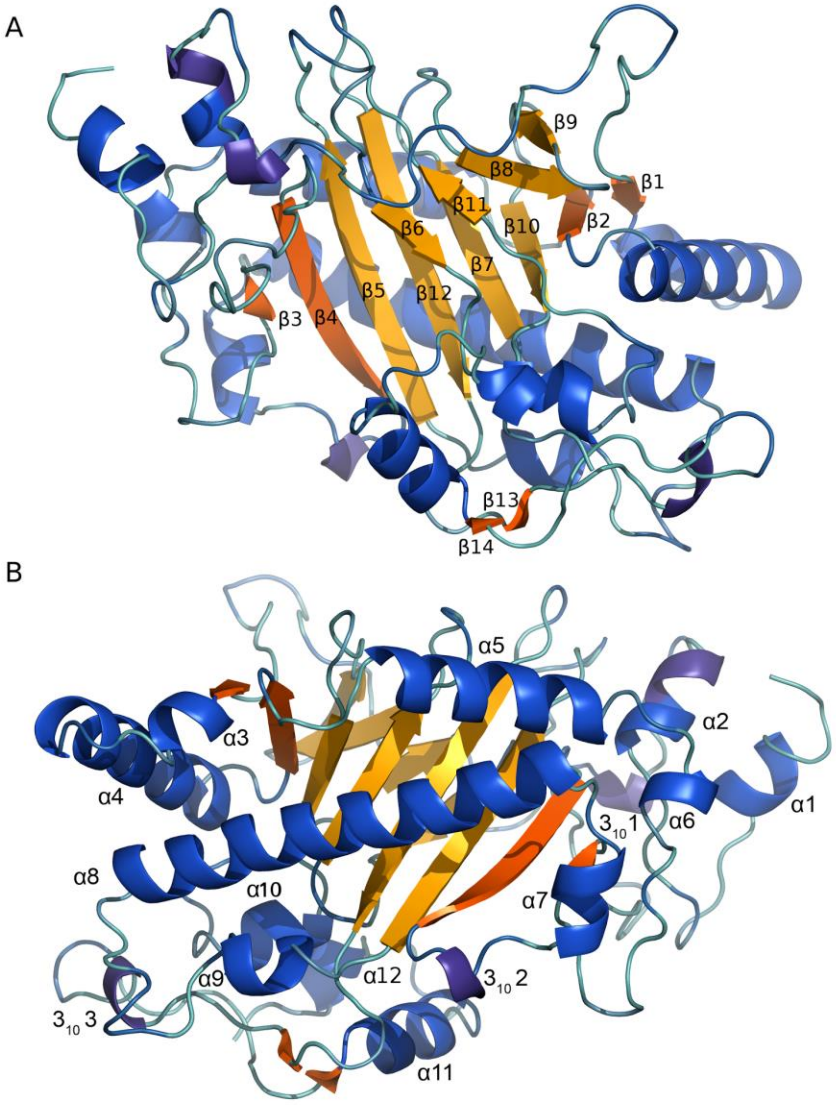

Figure 2. (A) Top and (B) bottom view of the overall structure of T6ODM (PDB: 5O7Y). Residues not included in the final model were rebuilt for clarity of visualization. Secondary structure assignment was performed with DSSP and presented with PyMOL using DSSP & Stride Plugin (Zhu, 2011; Kabsch and Sander, 1983; Schrödinger LLC, 2015; Touw et al., 2015). Helices (blue), β-strands (orange) are indicated and labeled. Conserved β-jellyroll fold of T6ODM is shown in yellow.

The active site

It was unclear whether the obtained structures contain nickel or iron ion bound in the active site. Results of preliminary ICP-OES analysis indicated high nickel and only residual iron content. The purification

protocol included Ni^{2+} affinity chromatography steps that sometimes, as a result of stripping nickel from the resin, may result in sample contamination (Niedzialkowska et al., 2017; Oswald et al., 1997). Nevertheless, results of previous structural studies on ODD indicate that the geometry of the active site and the protein, as a whole, is unaffected by Fe to Ni substitution (Horton et al., 2011). To determine which metal is present in the active site the data collection protocol by Handing et al. (Handing et al. in preparation) was employed. In short, the X-ray energy scan in the vicinity of absorption K-edge of nickel proved that nickel was present in the T6ODM:2OG crystals (**Figure S3**). To confirm that the nickel ion was present exactly in the active site, two diffraction data sets were collected at energies above and below the nickel absorption K-edge (**Figure S4**). For the two collected datasets we compared the ratio of peak values ($e/Å^3$) in an anomalous difference map of the metal ion present in the active site, to sulfur atoms in cysteine and methionine residues. At this wavelengths, the estimated ratios of anomalous scattering of atoms relevant to this structure are f''(Fe)/f''(S) = 5.80 (at 1.4831 Å), f''(Fe)/f''(S) = 5.80 (at 1.4902 Å), f''(Ni)/f''(S) = 7.51 (at 1.4831 Å) and f''(Ni)/f''(S) = 0.92 (at 1.4902 Å) as calculated with values estimated by the CCP4 CROSSEC (Cromer, 1983). The peak value ratios (calculated as ratios of the peak value of the ion in the active site, to the average value of sulfur atoms, with peak level above 4 r.m.s.d) calculated from the anomalous difference maps are 7.23 at 1.4831 Å and 0.92 at 1.4902 Å. Therefore, we have assigned a nickel ion in the active site.

Both 2OG and SIN are clearly visible in the electron density maps (**Figure 3. A-B**). The side chain of Arg-305 (highly conserved residue among ODD), located in strand β12, participates in stabilizing hydrogen bond interaction with the terminal carboxylate group of 2OG and SIN. Conserved amino acid residues: Asn-221, Tyr-223 (strand β5), Leu-235 (β6), Leu-247 (β7), Leu-256 (β8) and Ser-307 (β12) line up the co-substrate binding pocket of 2OG and SIN. When comparing this region of T6ODM to ANS and F6'H, it is noticeable that only the Leu-235 of T6ODM is not conserved and instead is replaced with Val residue.

In T6ODM:2OG complex the metal ion is coordinated by side chains of His-238, Asp-240, His-295, 1-carboxylate and 2-oxo groups of 2OG, and one water molecule. In the case of T6ODM in complex with SIN, the water molecule is absent in the metal coordination sphere; whilst SIN and two ethylene glycol molecules from either the crystallization buffer or cryoprotectant solution are modeled instead. The ethylene glycol molecule that is displaced by 2OG has a well resolved electron density but the identification of the other ethylene glycol molecule is more ambiguous. The electron density in this region was initially interpreted as two water molecules but, due to presence of the unresolved positive difference electron density between the water molecules, it was remodeled as ethylene glycol. In the T6ODM:2OG complex only two water molecules can be identified in this coordination site. In both structures we have observed elongated ambiguous electron densities that have been modeled tentatively as disordered PEG molecules.

The two determined structures (PDB: 5O7Y, 5O9W) can be superposed with a RMSD of 0.27 Å (as

calculated by PDBeFOLD) (Krissinel and Henrick, 2004). The replacement of succinate by 2OG results in a change of conformations or flexibility of Ile-148, Met-150 (β4) and Arg-219 (β5) that 'trap' 2OG in the active site (**Figure 3. C-D**). In this T6ODM:2OG model, Arg-219 adopts one major conformation, in which it is closer to the 2OG probably due to electrostatic interactions with the co-substrate, whereas there are two distinct Arg-219 conformations observed in the T6ODM:SIN complex. Due to change in Arg-219 orientation after 2OG binding, Ile-148 adopts a different conformation whereby it is not colliding with Arg-219. The shift of Arg-219 results in formation of additional space near Met-150, which can now help to facilitate binding of a hydrophobic moiety (modeled as PEG molecule) in a pocket formed between Val-128, Glu-129, Met-150 and Phe-152 side chains.

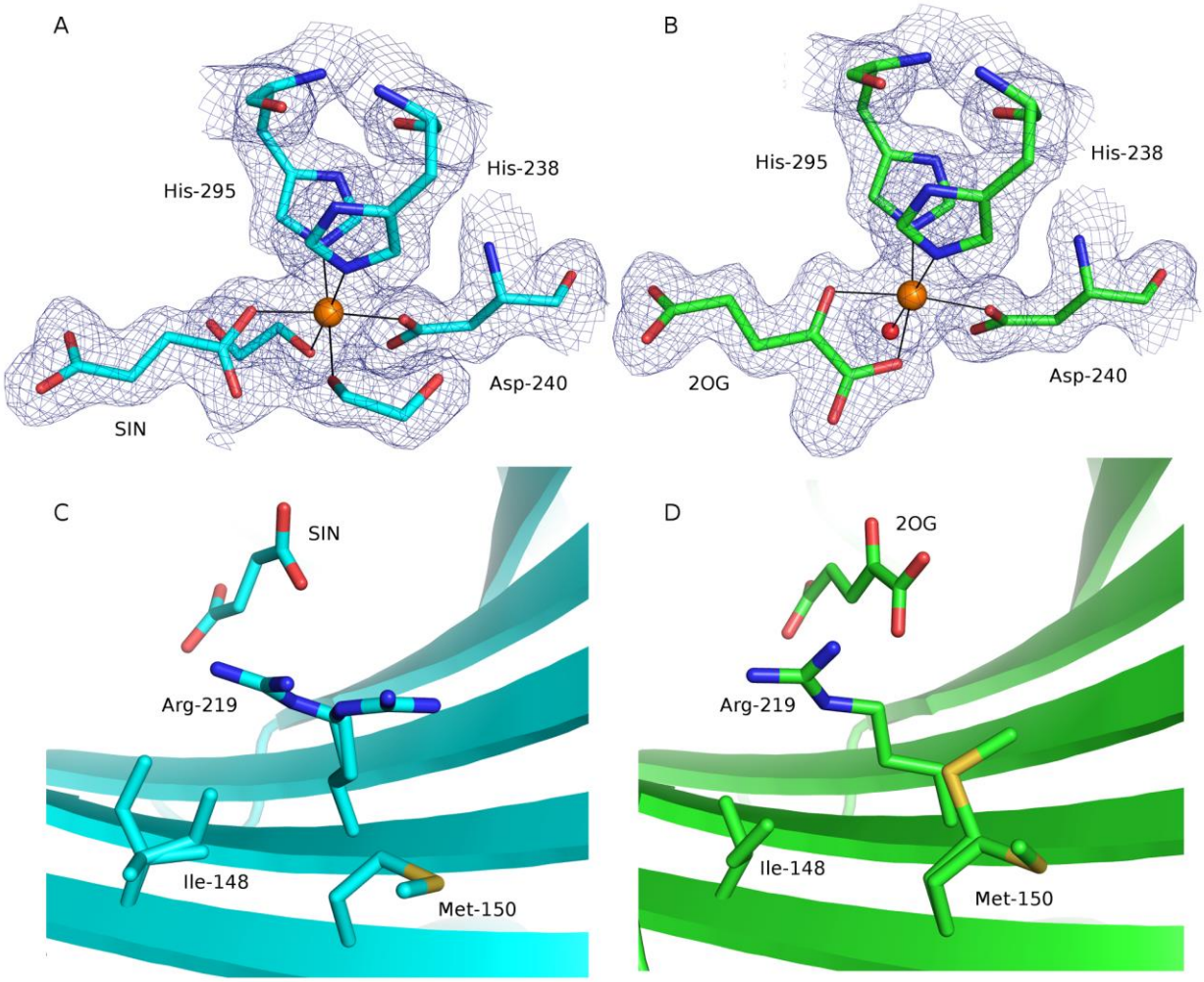

Figure 3. Active site of T6ODM complexed with (A) SIN and (B) 2OG shown with 2mF_o-DF_c electron density maps contoured at 1 r.m.s.d. Metal ion is shown as a sphere (orange), metal coordination is presented as solid line (black). Side chains arrangements with 2OG (C) or SIN (D) bound in the active site are presented. The models and the electron density maps, including omit maps, can be inspected

interactively using Molstack at http://molstack.bioreproducibility.org/project/view/OJM60NCMQF1VMKW0IBUF (T6ODM:2OG complex) and http://molstack.bioreproducibility.org/project/view/WU926J8HEKZP4ATBSW9B (T6ODM:SIN complex).

The C-terminal lid

T6ODM:SIN and T6ODM:2OG structures are generally well ordered with clear electron density maps. The structures were determined with an average B-factor of 29 $Å^2$ and 25 $Å^2$ for T6ODM:SIN and T6ODM:2OG, respectively. A few flexible regions that are characterized by relatively high B-factors can be identified: 36-53, 123-127, 200-204, 348-355 (including regions unmodeled due to uninterpretable electron density: residues 38-40, 46-51 for T6ODM:SIN and residues 37-40, 46-52 for T6ODM:2OG).

A flexible C-terminal loop along with surrounding α11 and α12 helices forms a lid over the active site. Previous mutagenesis studies of T6ODM and CODM demonstrated that this fragment is partly responsible for the specificity of these enzymes toward thebaine and codeine, respectively (Runguphan et al., 2012). Our structural comparison of the T6ODM with ANS shows that in ANS this lid is formed by an extended α-helix as compared to the helix-loop-helix motif in T6ODM (**Figure S2**). Therefore, former analyses that used comparative models of T6ODM and CODM based on ANS crystal structure may have not fully explored the importance of the C-terminal region for the substrate specificity of T6ODM and CODM enzymes. Since this fragment is supposed to play a significant role in the substrate selectivity, a detailed and reliable knowledge of its structure is essential for future studies.

The substrate binding cavity

The residues terminating the C-terminal loop (Thr-350 and Ser-357) are located 15 and 16 Å away from the nickel ion, respectively, whereas the most exposed ones (Leu-353, Asp-354) span to 22 Å from the cofactor ion. There are several possible explanations for how a lid that is located so far away from the catalytic site can influence substrate specificity. It could pre-orient the substrate during initial binding and be involved in a product-substrate exchange. The lid might close the entrance to the active site upon substrate binding in an induced-fit mechanisms. The hinge regions of the lid are involved in the crystal contacts (**Figure S5**), therefore it is possible that structure may be locked in this particular conformation. We speculate that this is a reason why, despite numerous attempts, we were unable to crystalize a T6ODM:morphine complex. On the other hand the related structure of the ANS does not demonstrate the conformational change in this part of the protein, upon substrate binding.

Furthermore, crystallographic studies of ANS revealed that it binds two molecules of substrate and a buffer molecule (Wilmouth et al., 2002) and since the binding pocket of T6ODM is also relatively spacious, it

is possible that it could accommodate two molecules of substrate as well (**Figure S6**). Additionally, it was suggested that presence of dihydroquercetin (substrate analog of ANS) in the distal binding site of ANS may hinder product release, which could inhibit enzyme activity at high substrate concentration (Welford et al., 2005). Substrate inhibition was observed for T6ODM for thebaine but not oripavine (Hagel and Facchini, 2010).

CONCLUSIONS

The results presented here describe the structure of T6ODM, an ODD enzyme involved in biosynthesis of pharmaceutically useful morphinan alkaloids. Although the general fold is usually maintained among ODD, binding of the substrates is accomplished by less conserved regions of the main core and surrounding loops. The T6ODM has a large substrate binding cavity that may allow the enzyme to bind large and diverse molecules - the feature that may contribute to T6ODM promiscuity. The size of the cavity is comparable to that of ANS and since the ANS cavity can accommodate multiple and various molecules at the same time, it is possible that the T6ODM substrate specificity and the kinetic characteristic changes when there are one or more compounds bound in the cavity. The previous studies show that the T6ODM activity with thebaine is affected by substrate inhibition - that observation is consistent with the hypothesis of multiple molecules bound in the cavity. However, substrate inhibition was not observed for oripavine. This poses an interesting question, whether the substrate specificity and regioselectivity are influenced by binding of the multiple molecules. We hope that the exact insight into the T6ODM structure we provided will help in future mutagenesis and computational studies aiming to understand the origins of substrate regioselectivity.

ACKNOWLEDGMENTS

This research project was supported by grant no. UMO-2014/14/E/NZ1/00053 from the National Science Centre, Poland, and iNEXT project no. 2477, funded by the Horizon 2020 programme of the European Union. The research was carried out with the equipment purchased thanks to the financial support of the European Regional Development Fund in the framework of the Polish Innovation Economy Operational Program (contract no. POIG.02.01.00-12-023/08). The synchrotron data was collected at beamline P13

operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank Dr Thomas Schneider for the assistance in using the beamline.

ACCESSION NUMBERS

The coordinate data files and structure factors have been deposited in the PDB under accession codes: 5O9W, 5O7Y.

SUPPORTING INFORMATION

Supplementary data associated with this article can be found at *DOI_of_Supporting_information*.

REFERENCES

Aik, W., McDonough, M.A., Thalhammer, A., Chowdhury, R., Schofield, C.J., 2012. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr. Opin. Struct. Biol. 22, 691–700.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. 66, 12–21.

Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S., Schneider, T.R., 2017. P13, the EMBL macromolecular crystallography beamline at the low-

emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. J. Synchrotron Radiat. 24, 323–332.

Cowtan, K., 2008. Fitting molecular fragments into electron density. Acta Crystallogr. Sect. D Biol. Crystallogr. 64, 83–89.

Cromer, D.T., 1983. Calculation of anomalous scattering factors at arbitrary wavelengths. J. Appl. Crystallogr. 16, 437–437.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501.

Eschenfeldt, W.H., Lucy, S., Millard, C.S., Joachimiak, A., Mark, I.D., 2009. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins, in: Methods in Molecular Biology (Clifton, N.J.). NIH Public Access, pp. 105–115.

Evans, P., 2006. Scaling and assessment of data quality. Acta Crystallogr. Sect. D Biol. Crystallogr. 62, 72–82.

Evans, P.R., 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 282–292.

Farrow, S.C., Facchini, P.J., 2013. Dioxygenases Catalyze O- Demethylation and O,O -Demethylenation with Widespread Roles in Benzylisoquinoline Alkaloid Metabolism in Opium Poppy. J. Biol. Chem. 288, 28997–29012.

French, G.S.,Wilson, K.S., 1978. On the treatment of negative intensity observations. Acta Crystallogr. A34, 517–525.

Grabowski, M., Langner, K.M., Cymborowski, M., Porebski, P.J., Sroka, P., Zheng, H., Cooper, D.R., Zimmerman, M.D., Elsliger, M.-A., Burley, S.K., Minor, W., 2016. A public database of macromolecular diffraction experiments. Acta Crystallogr. Sect. D Struct. Biol. 72, 1181–1193.

Hagel, J.M., Facchini, P.J., 2010. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. Nat. Chem. Biol. 6, 273–275.

Horton, J.R., Upadhyay, A.K., Hashimoto, H., Zhang, X., Cheng, X., 2011. Structural Basis for Human PHF2 Jumonji Domain Interaction with Metal Ions. J. Mol. Biol. 406, 1–8.

Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577–2637.

Kabsch, W., 2010. XDS. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 125–132.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2256–2268.

Krissinel, E., Henrick, K., 2007. Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. 372, 774–797.

Minor, W., Cymborowski, M., Otwinowski, Z., 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, 859–66.

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. REFMAC 5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 355–367.

Niedzialkowska, E., Mrugała, B., Rugor, A., Czub, M.P., Skotnicka, A., Cotelesage, J.J.H., George, G.N., Szaleniec, M., Minor, W., Lewiński, K., 2017. Optimization of overexpression of a chaperone protein

of steroid C25 dehydrogenase for biochemical and biophysical characterization. Protein Expr. Purif. 134, 47–62.

Oswald, T., Hornbostel, G., Rinas, U., Anspach, F.B., 1997. Purification of (His)6EcoRV [recombinant restriction endonuclease EcoRV fused to a (His)6 affinity domain] by metal-chelate affinity chromatography. Biotechnol. Appl. Biochem. 25 (Pt 2), 109–15.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–26.

Oxford Diffraction, 2006. CrysAlis Pro Oxford Diffraction Ltd, Version 1.171.36.20 (release 27-06- 2012 CrysAlis171.NET), Abingdon, England.

Porebski, P.J., Cymborowski, M., Pasenkiewicz-Gierula, M., Minor, W., 2016. Fitmunk: improving protein structures by accurate, automatic modeling of side-chain conformations. Acta Crystallogr. Sect. D, Struct. Biol. 72, 266–80.

Porebski, P.J., Sroka, P., Zheng, H., Cooper, D.R., Minor, W., 2017. Molstack-Interactive visualization tool for presentation, interpretation, and validation of macromolecules and electron density maps. Protein Sci.

Runguphan, W., Glenn, W.S., O'Connor, S.E., 2012. Redesign of a dioxygenase in morphine biosynthesis. Chem. Biol. 19, 674–678.

Sacksteder, K.A., Biery, B.J., Morrell, J.C., Goodman, B.K., Geisbrecht, B. V., Cox, R.P., Gould, S.J., Geraghty, M.T., 2000. Identification of the α-Aminoadipic Semialdehyde Synthase Gene, Which Is Defective in Familial Hyperlysinemia. Am. J. Hum. Genet. 66, 1736–1743.

Schrödinger LLC, 2015. The PyMOL Molecular Graphics System, Version 1.8, New York, USA.

Sun, X., Zhou, D., Kandavelu, P., Zhang, H., Yuan, Q., Wang, B.-C., Rose, J., Yan, Y., 2015. Structural Insights into Substrate Specificity of Feruloyl-CoA 6'-Hydroxylase from Arabidopsis thaliana. Sci. Rep. 5, 10355.

Touw, W.G., Baakman, C., Black, J., te Beek, T.A.H., Krieger, E., Joosten, R.P., Vriend, G., 2015. A series of PDB-related databanks for everyday needs. Nucleic Acids Res. 43, D364–D368.

Vagin, A., Teplyakov, A., 1997. MOLREP : an Automated Program for Molecular Replacement. J. Appl. Crystallogr. 30, 1022–1025.

Welford, R.W.D., Clifton, I.J., Turnbull, J.J., Wilson, S.C., Schofield, C.J., 2005. Structural and mechanistic studies on anthocyanidin synthase catalysed oxidation of flavanone substrates: the effect of C-2 stereochemistry on product selectivity and mechanism. Org. Biomol. Chem. 3, 3117–3126.

Wilmouth, R.C., Turnbull, J.J., Welford, R.W.D., Clifton, I.J., Prescott, A.G., Schofield, C.J., 2002. Structure and mechanism of anthocyanidin synthase from Arabidopsis thaliana. Structure 10, 93–103.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A.,

Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP 4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242.

Young, J.Y., Westbrook, J.D., Feng, Z., Sala, R., Peisach, E., Oldfield, T.J., Sen, S., Gutmanas, A., Armstrong, D.R., Berrisford, J.M., Chen, L., Chen, M., Di Costanzo, L., Dimitropoulos, D., Gao, G., Ghosh, S., Gore, S., Guranovic, V., Hendrickx, P.M.S., Hudson, B.P., Igarashi, R., Ikegawa, Y., Kobayashi, N., Lawson, C.L., Liang, Y., Mading, S., Mak, L., Mir, M.S., Mukhopadhyay, A., Patwardhan, A., Persikova, I., Rinaldi, L., Sanz-Garcia, E., Sekharan, M.R., Shao, C., Swaminathan, G.J., Tan, L., Ulrich, E.L., van Ginkel, G., Yamashita, R., Yang, H., Zhuravleva, M.A., Quesada, M., Kleywegt, G.J., Berman, H.M., Markley, J.L., Nakamura, H., Velankar, S., Burley, S.K., 2017. OneDep: Unified wwPDB System for Deposition, Biocuration, and Validation of Macromolecular Structures in the PDB Archive. Structure 25, 536–545.

Zheng, H., Chordia, M.D., Cooper, D.R., Chruszcz, M., Müller, P., Sheldrick, G.M., Minor, W., 2013. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat. Protoc. 9, 156–170.

Zheng, H., Cooper, D.R., Porebski, P.J., Shabalin, I.G., Handing, K.B., Minor, W., 2017. CheckMyMetal: a macromolecular metal-binding validation tool. Acta Crystallogr. Sect. D Biol. Crystallogr. 73, 223–233.

Zhu, H., 2011. DSSP & Stride plugin for PyMOL. Biotechnol. Cent. (BIOTEC), TU Dresden.